

Molekulare Mechanismen des hypoplastischen Linksherzsyndroms

Dissertation

zur Erlangung des Doktorgrades

der Mathematisch-Naturwissenschaftlichen Fakultät

der Christian-Albrechts-Universität zu Kiel

vorgelegt von

Kirstin Hoff

Kiel, 2016

Erster Gutachter: Prof. Dr. rer. nat. Hinrich Schulenburg

Zweiter Gutachter: Prof. Dr. med. Hans-Heiner Kramer

Tag der mündlichen Prüfung: 29.04.2016

Zum Druck genehmigt: 29.04.2016

gez. Prof. Dr. rer. nat. Wolfgang J. Duschl,
Dekan der Mathematisch-Naturwissenschaftlichen Fakultät
der Christian-Albrechts-Universität zu Kiel

1	EINLEITUNG	1
1.1	Anatomie und Funktion des Herzens	1
1.2	Herzentwicklung	4
1.3	Kongenitale Herzfehlbildungen.....	6
1.3.1	Arten kongenitaler Herzfehlbildungen und deren Pathophysiologie	7
1.3.2	Epidemiologie kongenitaler Herzfehlbildungen	9
1.4	Ätiologie kongenitaler Herzfehlbildungen	10
1.4.1	Genetische Faktoren bei kongenitalen Herzfehlbildungen.....	10
1.4.1.1	Chromosomale Anomalien und CHDs.....	11
1.4.1.2	Keimbahnmutationen und CHDs	12
1.4.1.3	Somatische Mutationen und CHDs	13
1.4.2	Epigenetische Aspekte kongenitaler Herzfehlbildungen	13
1.4.2.1	DNA-Methylierung als epigenetischer Regulationsmechanismus.....	15
1.5	Das hypoplastische Linksherzsyndrom.....	17
1.5.1	Anatomie und Pathophysiologie des HLHS	17
1.5.2	Ätiologie des HLHS	20
1.6	Hypothese und Zielsetzung.....	23
2	MATERIAL UND METHODEN	25
2.1	Material	25
2.1.1	Probenmaterial, Kontrollen und Zelllinien.....	25
2.1.1.1	Probenmaterial	26
2.1.1.2	Kontrollen und Zelllinien	35
2.1.2	Material für Laborarbeit	36
2.1.2.1	Primer	36
2.1.2.2	Chemikalien und Reagenzien.....	38
2.1.2.3	Puffer und Lösungen	39
2.1.2.4	Kits.....	40
2.1.2.5	Geräte und Zubehör	40
2.1.3	Software, R-Packages und Online-Datenbanken.....	42
2.1.3.1	Software und R-Packages.....	42
2.1.3.2	Online-Datenbanken und Online-Software-Anwendungen	44

2.2	Methoden	46
2.2.1	Molekulargenetische Analysen von DNA	46
2.2.1.1	Extraktion von genomischer DNA.....	46
2.2.1.2	Quantifizierung von DNA	48
2.2.1.3	Amplifikation von DNA-Fragmenten durch PCR	48
2.2.1.4	Agarose-Gelelektrophorese.....	50
2.2.1.5	Aufreinigung von PCR-Produkten	50
2.2.1.6	Sequenzierung von PCR-Produkten	51
2.2.1.7	Gesamt-Exom-Sequenzierung	52
2.2.1.8	SNP-Array Analyse.....	59
2.2.2	DNA-Methylierungsanalysen	60
2.2.2.1	Natriumbisulfit-Behandlung genomischer DNA.....	62
2.2.2.2	Array-basierte DNA-Methylierungsanalyse	63
2.2.2.3	PCR-basierte DNA-Methylierungsanalyse mittels Bisulfit-Pyrosequenzierung.....	68
2.2.2.4	PCR-basierte DNA-Methylierungsanalyse mittels MassARRAY-EpiTYPER-Technik	74
2.2.3	Molekulargenetische Analysen von RNA	80
2.2.3.1	Extraktion von RNA	80
2.2.3.2	Quantifizierung von RNA	81
2.2.3.3	Gesamt-Transkriptom-Sequenzierung	81
2.2.3.4	Reverse Transkription.....	82
2.2.3.5	Quantitative Real-Time PCR	83
2.2.4	Statistik	85
2.2.5	Bioinformatik.....	88
2.2.5.1	Bioinformatische Auswertung mittels R-Software	88
2.2.5.2	Bioinformatische Auswertung mittels Qlucore-Software	91
2.2.5.3	Funktionelle Anreicherungsanalysen und Gennetzwerk-Analysen	92
2.2.5.4	FastQC	93
3	ERGEBNISSE	94
3.1	Analysen auf Keimbahnebene	95
3.1.1	Identifizierung von Mutationen in <i>TGDS</i> als Beispiel einer syndrom-assoziierten kongenitalen Herzfehlbildung	95
3.1.2	Analysen nicht-syndromaler Herzfehlbildungen	98
3.1.2.1	Identifizierung einer <i>de novo</i> Mutation in <i>NR2F2</i> bei einem nicht-syndromalen HLHS-Fall	98

3.1.2.2	Analyse einer Familie mit zwei HLHS-Fällen: Identifizierung von <i>C15orf62</i> , <i>TNKS1BP1</i> und <i>PRF1</i> als potenziell krankheitsrelevante Kandidatengene nach autosomal dominantem Vererbungsmodell mit reduzierter Penetranz	100
3.1.2.3	SNP-Analyse eines nicht-syndromalen HLHS-Falls im familiären Kontext.....	103
3.2	Analysen auf somatischer Ebene	105
3.2.1	Genetische Analysen an Herzgewebe von Patienten mit HLHS.....	105
3.2.1.1	WES-Analyse an IAS-Gewebe von Patienten mit HLHS identifiziert 56 seltene bzw. bisher unbekannte <i>Loss-Of-Function</i> -Mutationen.....	105
3.2.1.2	Kombinierte Keimbahn- und somatische Analyse in einer Familie mit HLHS identifiziert <i>ANKRD11</i> und <i>BHLHA15</i> mit potenziellen herzwespezifischen Mutationen	117
3.2.2	DNA-Methylierungsanalysen an Herzgewebe von Patienten mit verschiedenen kongenitalen Herzfehlbildungen	120
3.2.2.1	Alterseffekte der DNA-Methylierung in IAS-Gewebeproben von Fällen mit HLHS und Rechtsherzfehlbildungen	125
3.2.2.2	Verschiedene Regionen des Herzens lassen sich durch differenzielle DNA-Methylierungsprofile unterscheiden	131
3.2.2.3	Nachweis einer differenziellen DNA-Methylierungs-Signatur in Atrium- und Ventrikelgewebe	139
3.2.2.4	Herzmuskelgewebe und Gefäßstrukturen des Herzens zeigen distinkte DNA-Methylierungs-Signaturen bei einem Patienten mit DCM.....	147
3.2.2.5	Identifizierung von 86 differenziell methylierten CpG-Loci in Herzgewebe von HLHS-Patienten im Vergleich zu fetalem Herzkontrollgewebe	152
3.2.3	Identifizierung von 11 differenziell exprimierten Genen bei HLHS im Vergleich zu Rechtsherzfehlbildungen	159
3.2.3.1	Qualitätskontrolle	160
3.2.3.2	Datenanalyse.....	161
3.2.3.3	Verifizierung.....	162
3.3	Integrative Analysen von Genom, Epigenom und Transkriptom	164
3.3.1	Identifizierung von 20 differenziell methylierten Genen mit zusätzlichen funktionellen Genvarianten bei Patienten mit CHDs	164
3.3.2	Identifizierung von <i>ACE2</i> und <i>SOSTDC1</i> mit einer differenziellen Genexpression und funktionellen Genvarianten bei Patienten mit HLHS	168
3.3.3	Identifizierung von <i>XYLT1</i> mit differenzieller DNA-Methylierung und aberranter Genexpression bei Patienten mit HLHS	171
3.3.4	Gesamtintegrative Analyse.....	172

4	DISKUSSION	173
4.1	Methodische Aspekte	175
4.1.1	Gewebeisolation und Herzgewebetyp	175
4.1.2	<i>Next Generation Sequencing</i> -Technologien	177
4.1.2.1	Gesamt-Exom-Sequenzierungstechnologie	177
4.1.2.2	RNA-Sequenzierung	180
4.1.3	Techniken der DNA-Methylierungs-Analyse	182
4.1.3.1	Natriumbisulfit-Konversion	182
4.1.3.2	450K-Array	183
4.1.3.3	Verifizierung von 450K-Ergebnissen mittels Bisulfit-Pyrosequenzierung	186
4.1.3.4	Anwendung der Epi typer-Technik zur Verifizierung von 450K-Ergebnissen	187
4.2	Biologische Aspekte	189
4.2.1	Identifizierung von Keimbahnmutationen bei syndromalen CHDs und nicht-syndromalen Fällen mit HLHS	189
4.2.1.1	Familien-Analyse von HLHS-Fällen: oligogene Vererbung mit reduzierter Penetranz und variabler Expressivität	191
4.2.2	Potenziell krankheitsrelevante Genvarianten im Herzgewebe von Patienten mit HLHS	194
4.2.2.1	Somatische Mutationen beim HLHS	199
4.2.3	Epigenetische Aspekte in der Analyse von Herzgewebe	201
4.2.3.1	Gewebespezifische DNA-Methylierungsprofile verschiedener Herzregionen	201
4.2.3.2	Altersspezifische und individuelle Effekte in der DNA-Methylierung	207
4.2.3.3	Potenzielle Kandidatengene mit differenzieller DNA-Methylierung im Herzgewebe von HLHS-Patienten im Vergleich zu fetalem Herzkontrollgewebe	209
4.2.4	Zusammenhänge auf Genom-, Epigenom- und Transkriptomenebene in der Entstehung des HLHS	211
4.3	Fazit	217
4.4	Ausblick	218
5	ZUSAMMENFASSUNG	220
6	SUMMARY	221
7	QUELLENVERZEICHNIS	224

8	ANHANG	264
9	DANKSAGUNG.....	281
10	LEBENS LAUF.....	282
11	PUBLIKATIONS LISTE.....	283
12	ERKLÄRUNG.....	286

Abkürzungsverzeichnis

1000G	1000 Genomes Project
450K-Array	Infinium HumanMethylation450 BeadChip Array
a	Jahre
A	Adenin (A = alt, Alterskategorie-Abkürzung in Tabelle 1)
aa	Aminosäuren
AF	Allelfrequenz
(m)AG	(mittleres) Alter bei Gewebeentnahme
ANOVA	<i>Analysis of Variance</i> (Varianzanalyse)
AA	Aortenstamm
Ao/AoV	Aortenklappe
AP	aortopulmonaler Shunt
AS	Aortenstenose
ASD	Atriumseptumdefekt
ATP	Adenosintriphosphat
AV(-Kanal)	Atrioventrikular(-kanal)
AVG-Beta	<i>average</i> Beta (mittlerer Beta-Wert)
AVSD	atrioventrikulärer Septumdefekt
b	<i>benign</i> (prädiktive Mutationskonsequenz)
BAV	bikuspide Aortenklappe
BLAT	<i>basic local alignment tool</i>
BMP	<i>bone morphogenic protein</i>
bp	Basenpaar
BSK-Sonde	Bisulfit-Konversionskontrollsonde
BSPS	Bisulfit-Pyrosequenzierung
C	Cytosin
°C	Grad Celsius
CoA	Koarktation der Aorta
CCG	Cologne Center for Genomics
CDD	Conserved Domains Database
CGI	CpG-Insel
cDNA	<i>complementary</i> DNA (komplementäre DNA)
CDS	<i>coding DNA sequence</i> (proteinkodierende Sequenz)
ChAS	Affymetrix® Chromosome Analysis Suite
CHD	kongenitale Herzfehlbildung (<i>congenital heart defect</i>)
Chr	Chromosom
cm	Zentimeter
CMS	Catell-Manzke-Syndrom
cNCC	kardiale Neuralleistenzellen (<i>cardiac neural crest cells</i>)
CNV	<i>copy number variation</i> (Kopienzahlvariation)

CpG	Cytosin-Guanin-Dinukleotid
Cq	<i>quantification cycle</i> (Quantifizierungs-Zyklus)
CRT	zyklisch-reversible Termination
CTL	Kontrolle
CV	<i>Reads Coverage</i>
d	Tag (<i>day</i>)
D	<i>probably damaging</i> (prädiktive Mutationskonsequenz)
Da	Dalton
DA	<i>Ductus arteriosus</i>
dA- <i>Tailing</i>	Inkorporation von Adenosinmonophosphat am 3'-Ende eines <i>blunt-end</i> DNA-Fragments
dATP α S	Desoxyadenosin-Alpha-Thiotriphosphat
DB	Datenbank
dbSNP	dbSNP (Short Genetic Variations)
DCM	Dilatative Kardiomyopathie
dCTP	Desoxycytidintriphosphat
DD	Entwicklungsstörung (<i>developmental disorder</i>)
ddH ₂ O	doppelt destilliertes Wasser
(d)dNTP	(Di)desoxynukleosidtriphosphat
del	Deletion
<i>Detection Pval</i>	Detektions-p-Wert
dGTP	Desoxyguanosintriphosphat
DGV	Database of Genomic Variants
DHS	<i>DNase I hypersensitive site</i>
DM	DNA-Methylierungsanalyse (Abkürzung in Tabelle 1)
DM-CpG-Locus	differenziell methylierter CpG-Locus
DMR	differenziell methylierte Region
DMSO	Dimethylsulfoxid
DNA	Desoxyribonukleinsäure
DNMT	DNA-Methyltransferase
DNP	2,4-Dinitrophenol
DPC	<i>days post coitum</i>
dsDNA	Doppelstrang-DNA
DSMZ	Deutsche Sammlung von Mikroorganismen und Zellkulturen
dTTP	Desoxythymidintriphosphat
dup	Duplikation
E	Primereffizienz (qPCR), bzw. Gesamt-Exom-Sequenzierung (Abkürzung in Tabelle 1)
EDTA	Ethylendiamintetraessigsäure
EFE	Endokardfibroelastose
ELM	Eukaryotic Linear Motif resource, Functional Sites in Protein
EndMT	endothelial-mesenchymale Transition
EPCC	European Paediatric Cardiac Code
EPDC	Epikard-stämmige Zellen (<i>epicardium-derived cells</i>)

EpiTyper-Technik	MassARRAY® EpiTYPER® Technik
ESP	NHLBI Exome Sequencing Project
EVS	NHLBI Exome Sequencing Project, Exonic Variant Server
ExAC	Exome Aggregation Consortium
EZM	Extrazelluläre Matrix
F/fw	Vorwärts-Primer
FASD	fetale Alkohol-Spektrum-Störungen
FDR	<i>False Discovery Rate</i>
FE	arbiträre Fluoreszenzeinheiten (<i>Quality Control Bar Plot</i>)
FGF	<i>fibroblast growth factor</i>
FHF	primäres Herzfeld (<i>first heart field</i>)
FO	<i>Foramen ovale</i>
fr	frisch
FS	Frameshift
g	Einheit Gramm (im Kontext mit der Nennung von Zentrifugen g = Erdbeschleunigung)
G	Guanin
gDNA	genomische DNA
GH	Gesamtherz
GL	Glenn-Anastomose
GO	<i>Gene Ontology</i>
GOF	<i>Gain-of-Function</i>
GRCh37/hg19	Genome Reference Consortium (GRC) released "build 37" des humanen Genoms (hg19)
GS	Genome Studio
h	<i>hour</i> (Stunde)
H	Hemifontan-Operation
Herz-Genliste	Liste von Genen mit bekanntem kardialen Phänotyp in Maus/Human
HGNC	HUGO Gene Nomenclature Committee
H ₂ O	Wasser
HGVS	Human Genome Variation Society
HLHS	Hypoplastisches Linksherzsyndrom
hMC	<i>homogeneous</i> MassCLEAVE™
HPRT1	Hypoxanthin-Phosphoribosyl-Transferase 1
HTx	Herztransplantation
HTML	<i>Hypertext Markup Language</i>
HZV	Herzzeitvolumen
IAS	interatriales Septum
ID	Identifizierung
IDAT	<i>Intensity Data files</i> (Rohwerte 450K-Array)
IKMB	Institut für Klinische Molekularbiologie
INDEL	Kombination aus Insertion und Deletion
iPSC	induzierbare pluripotente Stammzelle
IVC	<i>Vena cava inferior</i> (untere Hohlvene)

J	jung/junge Alterskategorie (Abkürzung in Tabelle 1)
kbp	Kilo-Basenpaare
l	Liter
LA	linkes Atrium
LCR	<i>Loci Call Rate</i>
LOF	<i>loss-of-function</i> (Funktionsverlustmutation)
LOH	<i>loss-of-heterozygosity</i>
LS-CHD	linksseitige kongenitale Herzfehlbildungen
LV	linker Ventrikel
LV-An/Lvan	linksventrikuläres Aneurysma
LVNC	linksventrikuläre <i>non-compaction</i> Kardiomyopathie
LVOT	linksventrikulärer Ausflusstrakt
LVOTO	linksventrikuläre Ausflusstrakt-Obstruktion
m	männlich
M	Mol, bzw. mittel/mittlere Alterskategorie (Abkürzung in Tabelle 1)
MA/AA	Mitralatresie/Aortenatresie
MAF	<i>mutated allele frequency</i> (mutierte Allelfrequenz)
MALDI-TOF-MS	Matrix-unterstützte Laser-Desorption/Ionisation Flugzeit-Massenspektrometrie
Mapinfo	chromosomale Position nach GRCh37/hg19
Mb	Mega-Basenpaare
MBD	<i>Methyl-CpG Binding Domain</i>
5-mC	methyliertes Cytosin
MgCl ₂	Magnesiumchlorid
MGI	Mouse Genome Informatics
min	Minute
Miss.	Missense-Mutation
MK	Methylierte Kontrolle
mm	Millimeter
mRNA	<i>messenger</i> Ribonukleinsäure
MS/AA	Mitralstenose/Aortenatresie
MS/AS	Mitralstenose/Aortenstenose
MV	Mitralklappe
MVP	Mitralklappen-Prolaps
MW	Mittelwert
N	Anzahl bzw. N = Norwood-Operation (Abkürzung in Tabelle 1)
NA	keine Angabe/kein Messwert vorhanden
NaCl	Natriumchlorid
NCBI	National Center for Biotechnology Information
NGS	<i>Next Generation Sequencing</i>
n.c.	nicht gecouvert
ng	Nanogramm
n.m.	nicht möglich

Nomral-GS-Tabelle	normalisierte GS-gefilterte CpG-Loci-Tabelle
N_Shore	<i>North Shore</i> einer CpG-Insel
n.p.	nicht publiziert
n.t.	nicht getestet
Oligo-dT	Oligo-dT-Nukleotide (15 – 25 Desoxythymidine), welche komplementär zum Poly-A-Ende der eukaryotischen mRNA sind
OMIM	Online Mendelian Inheritance in Man
OP	Operation
OR	<i>odds ratio</i>
ORF	<i>open reading frame</i> (offenes Leseraster)
P	<i>possibly damaging</i> (prädiktive Mutationskonsequenz)
PA	Pulmonalarterie
PAt	Pulmonalatresie
pB	peripheres Blut
PCA	<i>Principal Component Analysis</i>
PCR	<i>polymerase chain reaction</i> (Polymerase-Kettenreaktion)
p.d.	<i>probably damaging</i> (prädiktive Mutationskonsequenz)
PDA	persistierender <i>Ductus arteriosus</i>
PE	Proepikardium
p.k.	potenziell krankheitsassoziiert
pm	<i>post mortem</i>
pmol	Pikomol
PPH	Polymorphism Phenotyping v2 (Polyphen-2)
PS	Pulmonalstenose
QC	Qualitätskontrolle
PV	Pulmonalvene
qPCR	quantitative <i>Real-Time</i> PCR
q-Wert	statistische Einheit der FDR
r	Korrelationskoeffizient
R ²	Bestimmtheitsmaß
R/rv	Rückwärts-Primer
R1/R2	Gesamt-Transkriptom-Sequenzierung Analyse 1 bzw. 2 (Abkürzung in Tabelle 1)
RA	rechtes Atrium
RDMR	reprogrammierungsspezifische differenziell methylierte Region
RE	Restriktionsenzym
RF	Rechtsherzfehlbildungen
RNA	Ribonukleinsäure
RNA-Sequenzierung	Gesamt-Transkriptom-Sequenzierung
RP	Rückwärts-Primer
RPS	RVIS-Perzentil-Score
rs	Referenz-SNP
RT	reverse Transkriptase

RV	rechter Ventrikel
RVIS	<i>Residual Variation Intolerance Score</i>
s	Sekunde
S	Sequenzier-Primer
SAM	S-Adenosyl-Methionin
Sanger-Sequenzierung	Sequenzierung von PCR-Produkten nach Sanger-Methode
SD	Standardabweichung
SDR	<i>short-chain</i> Dehydrogenase/Reduktase
SHF	sekundäres Herzfeld (<i>second heart field</i>)
S_Shore	<i>South Shore</i> einer CpG-Insel
siRNA	small interfering RNA
SM	<i>silence methylated</i>
SN	<i>silence nonmethylated</i>
SNP	<i>single nucleotide polymorphism</i> (Einzel-Nukleotid-Polymorphismen)
SNP-Array	Genome-Wide Human SNP Array 6.0
SNV	single-nucleotide variant
ssDNA	<i>single-stranded</i> (Einzelstrang-) DNA
SSW	Schwangerschaftswoche
Sv	<i>Sinus venosus</i>
SVC	<i>Vena cava superior</i> (obere Hohlvene)
SWAN	<i>Subset-quantile Within Array Normalization</i>
T	Thymin
TA	Trikuspidalatresie
Tag	Sequenz-Anhang
TargetID	Illumina CpG Identifikationsnummer
TCPC	totale kavopulmonale Anastomose
TBE	TRIS-Borat-EDTA Puffer
TF	Transkriptionsfaktor
TFBS	Transkriptionsfaktorbindestelle
TGA	Transposition der großen Arterien
TOF	Fallot'sche Tetralogie
Tm	Primer Schmelztemperatur
TRIS	Tris(hydroxymethyl)-aminomethan
TriV/TV	Trikuspidalklappe
UCSC	<i>University of California Santa Cruz (UCSC) genome assemblies</i>
UKSH	Universitätsklinikum Schleswig-Holstein
U/min	Umdrehung pro Minute
UTR	<i>untranslated region</i>
V	Volt
VA	visuelle <i>Reads</i> -Analyse
VP	Vorwärtsprimer
VSD	Ventrikelseptumdefekt

w	weiblich
W	Woche
WebGestalt	WEB-based GEne SeT AnaLysis Toolkit
WES	<i>Whole Exome Sequencing</i> (Gesamt-Exom-Sequenzierung)
WGA	<i>Whole Genome Amplification</i> (Gesamt-Genom-Amplifikation)
wT-Test	<i>Welch Two Sample</i> T-Test
$\Delta\beta$	Delta.Beta (Methylierungsdifferenz der AVG-Beta-Werte)
μg	Mikrogramm
μl	Mikroliter
λ	Lamda (Wellenlänge)

1 Einleitung

Das kardiovaskuläre System wurde erstmalig in der Zeit des antiken Griechenlands durch Galen [1] (129 - 216 AD) beschrieben. Ein tiefergehendes Verständnis über die Physiologie, Anatomie und Pathologie des Herzens wurde dann später durch Ibn al-Nafis [2] und William Harvey [3] erarbeitet [4]. Das menschliche Herz ist ein muskuläres Hohlorgan, welches durch rhythmische Kontraktion die Bewegung des Bluts durch den Kreislauf ermöglicht [5]. Es gehört zu den ersten während der Embryonalentwicklung angelegten Organen und sichert die Versorgung dieser mit Sauerstoff und Nährstoffen [6]. Als zentrales Organ des Blutkreislaufs stellt es den Motor des gesamten Organismus dar und ist von essenzieller Bedeutung für das Erhalten lebenswichtiger Funktionen des Körpers [7]. Liegt eine Fehlbildung des Herzens vor, so kommt es zu Funktionseinschränkungen des Herz-Kreislaufs- bzw. des Herz-Lungen-Systems [8]. Das Spektrum angeborener Herzfehlbildungen reicht von einfachen Herzfehlbildungen mit geringen Funktionsbeeinträchtigungen, bis zu schwerwiegenden kardialen Anomalien, die unbehandelt zum Tode führen [9]. Zu letztgenannten zählt das hypoplastische Linksherzsyndrom (HLHS). Die Erforschung seiner molekularen Mechanismen ist Thematik der vorliegenden Arbeit.

1.1 Anatomie und Funktion des Herzens

Das Herz liegt links versetzt hinter dem Brustbein mit seitlich angrenzender linker und rechter Lunge, getrennt durch parietale und viszerale Pleura [5]. Es ist von bindegewebshaltigem Perikard (Herzbeutel) umschlossen [10]. Am Abgang der großen Gefäße schlägt die innerste Schicht des Perikards in das Epikard (äußeres Bindegewebe des Herzens) um, welches fest mit dem Herzen verwachsen ist [10]. In einer Fettgewebeschicht unter dem Epikard verlaufen die Herzkranzgefäße, die sich in die darunterliegende Muskelschicht, dem Myokard, erstrecken und die Blutversorgung des Herzmuskels sichern [7]. Das Myokard besteht aus spezialisiertem Muskelgewebe (Kardiomyozyten), das nur im Herzen vorkommt [7]. Die innerste Schicht bildet das bindegewebshaltige Endokard mit einer Epithelzellschicht zum Herzinnenraum [7].

Das menschliche Herz besteht aus einer rechten und linken Herzhälfte, die jeweils aus einer Kammer (*Ventriculus cordis*, Herz-Ventrikel) und einem Vorhof (Atrium) besteht [5] (Abbildung 1). Eine Vorhofscheidewand (*Septum interatriale*, interatriales Septum)

trennt die Atrien voneinander und eine Kammerscheidewand (Septum *interventriculare*, Ventrikelseptum) separiert die Ventrikel [5]. Die Septen bestehen jeweils aus einem Endokardanteil, wobei das Ventrikelseptum zusätzlich auch einen muskulären Anteil enthält [7].

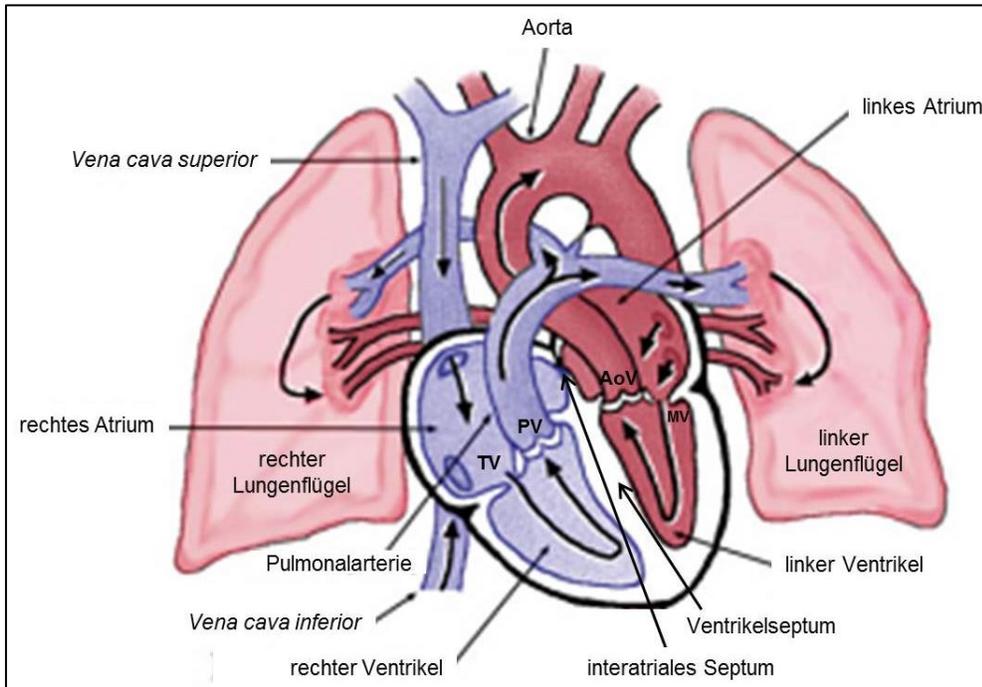


Abbildung 1 Anatomie des menschlichen Herzes. Dargestellt ist das Herz mit dem umliegenden Gefäßsystem und die Zu- und Abgänge zu den Lungenflügeln. Das desoxygenierte Blut (stammt aus dem Körperkreislauf) ist in blau dargestellt und oxygeniertes Blut (stammt aus dem Lungenkreislauf) in rot. Die Blutflussrichtung ist durch Pfeile verdeutlicht. Desoxygeniertes Blut gelangt hierbei über die obere und untere Hohlvene (*Vena cava superior* und *inferior*) in den rechten Vorhof, über die Trikuspidalklappe (TV) in den rechten Ventrikel und von dort über die Pulmonalklappe (PV) und Pulmonalarterie in den Lungenkreislauf. Das dort mit Sauerstoff angereicherte Blut gelangt anschließend in den linken Vorhof, über die Mitralklappe (MV) in den linken Ventrikel und über die Aortenklappe (AoV) und Aorta in den Körperkreislauf. (Abbildung aus laizzo (2010) [11], modifiziert)

Die physiologischen Gegebenheiten des Herzens werden in den folgenden zwei Absätzen nach dem Lehrbuch von Katz (2010) „*Physiology of the Heart*“ [7] beschrieben.

Ein gerichteter Blutfluss durch die Herzzräume erfolgt mithilfe von Herzklappen (Abbildung 1), die als „Rückschlagventile“ fungieren und an einer Bindegewebsplatte, dem Herzskelett, aufgehängt sind. Zwischen Atrien und Ventrikel befindet sich jeweils eine Segelklappe: (1) die 2-zipfelige Mitralklappe zwischen linkem Atrium und linkem Ventrikel, bestehend aus zwei von glattem Endothel überzogenen Einzelsegeln, und (2) die 3-zipfelige Trikuspidalklappe zwischen rechtem Atrium und rechtem Ventrikel, die aus drei Klappensegeln besteht. Die großflächigen Segelklappen sind durch Sehnenfäden (*Chordae tendineae*) an der Herzinnenwand über Papillarmuskeln befestigt. Durch Taschenklappen sind die Ventrikel mit den großen abgehenden Arterien verbunden. Die Aortenklappe trennt den linken Ventrikel von der abgehenden Aorta *ascendens*, während die Pulmonalklappe den rechten Ventrikel von dem Gefäßstamm der

Lungenarterien (*Truncus pulmonalis*) trennt. Die Taschenklappen bestehen aus drei etwa gleichgroßen Taschen und verhindern einen Rückfluss des Blutes aus den Arterien in die Ventrikel. ([7] S. 3 ff)

Im gesunden Herzen sind die linke und rechte Herzhälfte vollständig voneinander getrennt und Gleiches trifft für den Lungen- und Körperkreislauf zu. Sauerstoffarmes Blut wird mithilfe der oberen und unteren Hohlvene (*Vena cava superior* und *inferior*) von den Organen zum Herzen transportiert. Es gelangt dann über den rechten Vorhof durch die Trikuspidalklappe in den rechten Ventrikel, welcher das Blut durch die Pulmonalklappe dem Lungenkreislauf zuführt (Abbildung 1). Das in der Lunge mit Sauerstoff angereicherte Blut fließt anschließend über Lungenvenen in den linken Vorhof und durch die Mitralklappe in den linken Ventrikel. Durch den sogenannten linksventrikulären Ausflusstrakt (LVOT) gelangt das sauerstoffreiche Blut über die Aortenklappe und die Aorta in den Körperkreislauf. Im Vergleich zu dem Lungenkreislauf ist der Gesamtgefäßwiderstand und Blutdruck im Körperkreislauf deutlich höher, was sich in der stärkeren Wanddicke des linken Ventrikels verglichen mit dem rechten widerspiegelt. ([7] S. 297 ff)

Die **fetale Herzzirkulation** unterscheidet sich von der postnatalen Zirkulation, da die Lunge noch nicht in Funktion tritt [12]. In diesem Abschnitt wird die fetale Zirkulation nach dem Lehrbuch von Apitz (2013) „Pädiatrische Kardiologie: Erkrankungen des Herzens bei Neugeborenen, Säuglingen, Kindern und Heranwachsenden“ [13] beschrieben. Oxygeniertes und nährstoffreiches Blut der maternalen Zirkulation aus der Plazenta erreicht über die Hohlvenen den rechten Vorhof des Fetus. Durch zwei natürliche fetale Shunts (Verbindungen, die den Blutfluss zwischen normalerweise getrennten Gefäßen ermöglichen), dem *Ductus arteriosus* (DA) und dem *Foramen ovale* (FO), wird ein Übertritt dieses Mischbluts in die linke Herzhälfte ermöglicht. Ein Großteil des Bluts gelangt über das FO (Durchtritt durch das interatriale Septum) direkt von dem rechten Vorhof in den linken und wird von dort über den linken Ventrikel und die Aorta in den Körperkreislauf gepumpt. Ein kleiner Anteil des Bluts gelangt über die Trikuspidalklappe und den rechten Ventrikel in den *Truncus pulmonalis*. Aufgrund des hohen Gefäßwiderstands der noch nicht funktionellen Lunge wird das Blut durch den DA, der Gefäßverbindung zwischen dem Hauptstamm der Lungenarterie und der Aorta *ascendens*, in den Körperkreislauf befördert. Nach der Geburt verringert sich der Lungengefäßwiderstand durch die Ventilation der Lunge mit einhergehender Vasodilation der Lungenarterien und innerhalb von 10 - 15 Stunden verschließen sich das FO und der DA, sodass zwei parallele getrennte Kreisläufe, Körperkreislauf und Lungenkreislauf, resultieren. ([13] S. 35 ff.)

1.2 Herzentwicklung

Bereits in der dritten Woche der Embryonalentwicklung beginnt die Ausbildung des menschlichen Herzens [6]. Dieser Prozess lässt sich in vier Phasen unterteilen (Abbildung 2 A): (1) die Bildung der kardialen *Crescent* (hufeisenförmige Struktur bestehend aus kardialen Vorläuferzellen), (2) die Ausbildung des primitiven Herzschauchs, (3) *Looping* des Herzschauchs und (4) Bildung der Septen, Gefäße und der Herzklappen [14].

Erste kardiale Strukturen beginnen sich in der dritten Woche der Embryonalentwicklung als **kardiale *Crescent*** auszubilden (15 DPC (*days post coitum*)) [14] (Abbildung 2 A). Mithilfe kardialer TF-Gene, wie *GATA4*, *NKX2.5* und *TBX5*, deren Expression räumlich und zeitlich durch Wnt-, FGF (*fibroblast growth factor*)- und BMP (*bone morphogenic protein*)-Signalgebung kontrolliert wird, werden die Herzfelder etabliert [15 - 19]. Das primäre Herzfeld (FHF), besteht aus mesodermalen kardialen Vorläuferzellen, die die kardiale *Crescent* bilden [20]. Das sekundäre Herzfeld (SHF) stellt eine Subpopulation aus (noch undifferenziertem) kardialen Mesoderm dar [16]. Diese Zellpopulation liegt entlang der kardialen *Crescent* und wird später graduell zu dem sich bildenden Herzschauch hinzugefügt [16]. Basierend auf FHF-Zellen beginnt sich anschließend der **primitive Herzschauch** zu bilden [16] (Abbildung 2 A). Notch-, BMP- und Wnt-Signalgebung ermöglichen die Differenzierung der kardialen Vorläuferzellen, sodass verschiedene Zellsubpopulationen des Herzens entstehen [21]. Hierzu gehören u.a. Kardiomyozyten, Endothelzellen, Fibroblasten und glatte Muskelzellen [21]. Das Wachstum des Herzschauchs wird durch ein graduelles Einwandern von SHF-Zellen unterstützt [16, 22]. Der sich bildende Herzschauch ermöglicht nun erstmals einen unidirektionalen Blufuss durch peristaltik-ähnliche Bewegung [23]. Der erste Herzschlag setzt ab 21 DPC ein [14].

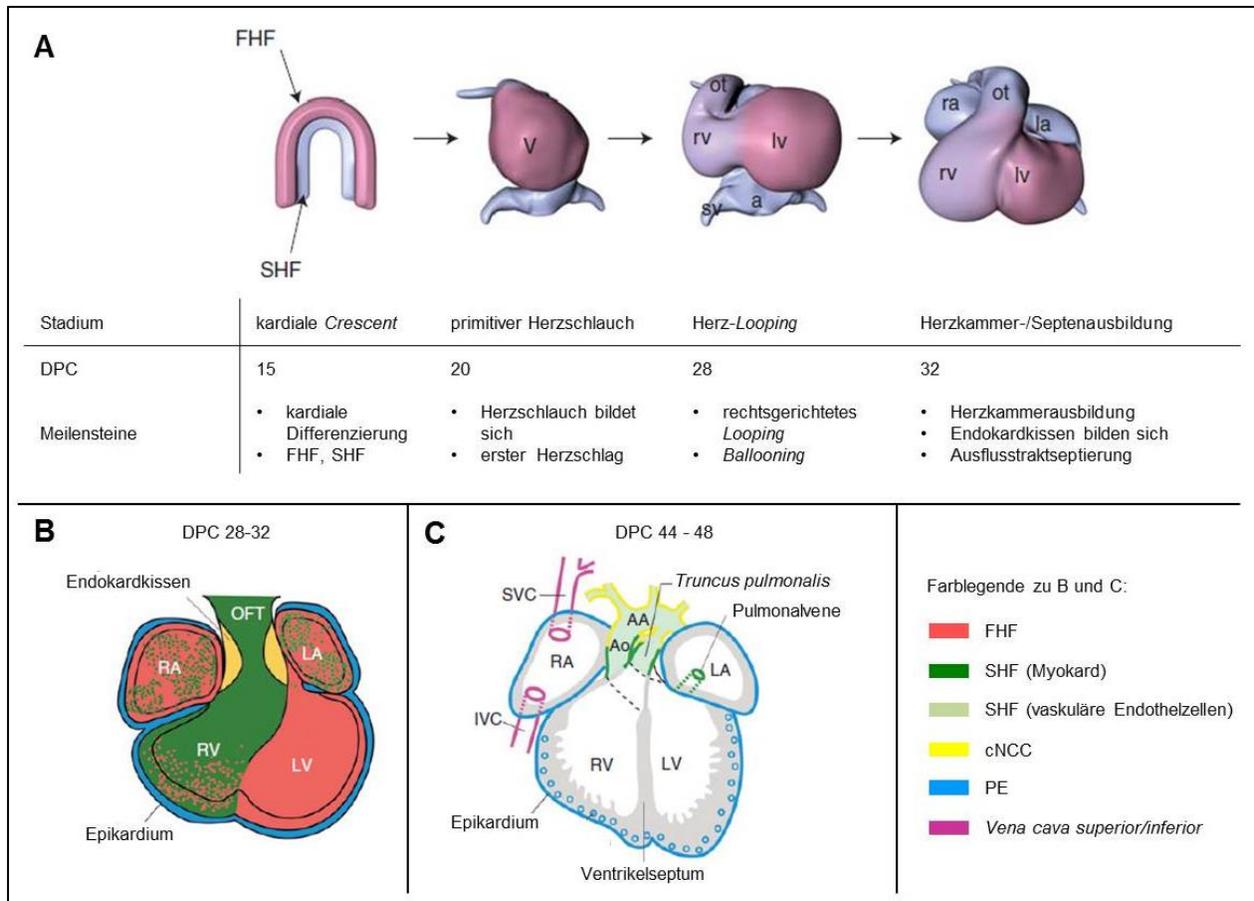


Abbildung 2 Schematische Darstellung der Herzentwicklung. A: Meilensteine der Herzentwicklung mit Angabe zu DPC (*days post coitum*)-Zeitraum der vier Stadien in der Herzentwicklung. Zellen des FHF's (primäres Herzfeld) sind in rot, die des SHF's (sekundäres Herzfeld) in blau dargestellt. B: Darstellung des Herzens im Zeitraum DPC 28 – 32 nach erfolgtem Herz-*Looping* und Beginn der Herzkammerausbildung. C: Darstellung des Herzens nach der Septierung (DPC 44 – 48). Farblegende: FHF und Derivate in rot, SHF und dessen Derivate in dunkelgrün (Myokard) und blassgrün (vaskuläre Endothelzellen), cNCCs (kardiale Neuralleistenzellen) in gelb (vaskuläre glatte Muskelzellen und Endokardkissen) und Zellen des PE's (Proepikardium) in blau. Abkürzungen: v: Ventrikel, a: Atrium, RA: rechtes Atrium, LA: linkes Atrium, RV: rechter Ventrikel, LV: linker Ventrikel, OFT/ot: Ausflusstrakt, sv: *Sinus venosus* (venöser Pol), AA: Aortenstamm, Ao: Aorta, IVC: *Vena cava inferior*, SVC: *Vena cava superior*. (A: Abbildung aus Bruneau (2008) [24], modifiziert; B und C: Abbildung aus Vincent and Buckingham (2010) [21], modifiziert)

Während des anschließenden **Loopings** beginnt die Ausbildung der kardialen Kammern. Durch das sogenannte **Ballooning** weitet sich hierbei der einzelne Ventrikel des Herzschlauchs seitlich auf [25, 26] (Abbildung 2 A). Das Einwandern myokardialer Zellen des SHF's an beiden Polen des tubulären Herzens ermöglicht die Bildung des Ausflusstrakts [27], sowie, in Kombination mit FHF-Zellen, die Ausbildung des rechten Ventrikels und später der Atrien [28 - 30]. Während des rechts-gerichteten *Looping*-Prozesses bilden sich auf Höhe des Atrioventrikularkanal (AV-Kanal, Bereich des späteren atrialem und ventrikulärem Septums) und des Ausflusstrakts sogenannte Endokardkissen [21]. Diese bestehen aus einer Subpopulation mesenchymaler Zellen und sind später für die Ausbildung der Septen und Herzklappen von Bedeutung [31 - 33].

Am venösen Pol des Herzschlauchs erfolgt die Entwicklung des Proepikardiums (PE) (23 - 26 DPC), welches das Epikard und Epikard-stämmige Zellen (*epicardium-derived cells* (EPDC)) bildet [34]. Ab 26 - 30 DPC beginnt die **Ausbildung von Septen** (Abbildung 2 A) unter der Beteiligung kardialer Neuralleistenzellen (cNCCs) sowie der Endokardissen, sodass eine Septierung von Vorhöfen, Ventrikel und kardialem Ausflusstrakt resultiert [14]. Parallel beginnt nun auch die **Gefäßentwicklung** mit der Ausbildung des aortalen Stamms und des *Truncus pulmonalis*, unterstützt durch die Migration von cNCCs als mesenchymale Vorläuferzellen in den kardialen Ausflusstrakt [33, 35]. Die **semilunaren Klappen**, Aortenklappe und Pulmonalklappe, entwickeln sich ab 42 – 44 DPC. Anschließend beginnen sich die **atrioventrikulären Klappen**, Mitralklappe und Trikuspidalklappe, zu bilden (44 - 48 DPC) [14]. Resultat der beschriebenen Entwicklungsphasen ist ein vierkammriges Herz (DPC 60) [14].

Insgesamt stellt die Herzentwicklung einen komplexen, vom Insekt bis hin zu Vertebraten konservierten Prozess dar, der auf einer räumlich und zeitlich fein abgestimmten Differenzierung kardialer Vorläuferzellen basiert [14]. Die Ausbildung eines funktionellen Herzens wird hierbei durch ein koordiniertes Zusammenspiel genetischer und epigenetischer Faktoren [36 - 39], flussdynamischer Stimuli und intra-/interzellulärer Signalgebung reguliert [40 - 42]. Während dieser Phase der Kardiogenese können unterschiedliche Einflüsse zu einer fehlerhaften Entwicklung und Schädigung des Herzens führen, die sich in der Ausbildung kongenitaler Herzfehlbildungen wie dem hypoplastischen Linksherzsyndrom widerspiegeln [43]. Insgesamt tragen Aberrationen in der Herzentwicklung zu einer signifikant erhöhten perinatalen Mortalität bei [44, 45].

1.3 Kongenitale Herzfehlbildungen

Kongenitale Herzfehlbildungen (*congenital heart defects*, CHDs) lassen sich nach der Definition von Mitchell et al. (1971) als „*gross structural abnormality of the heart or intrathoracic great vessels that is actually or potentially of functional significance*“ [46] definieren. CHDs umfassen somit angeborene strukturelle Anomalien des Herzens oder angrenzender, intrathorakaler großer Gefäße, die je nach Art zu unterschiedlich schwerwiegenden Funktionseinschränkungen des Herz-Kreislaufs- bzw. des Herz-Lungen-Systems führen [46].

CHDs stellen eine der häufigsten menschlichen Fehlbildungen dar und werden als wesentliche Herausforderung der Gesundheitsversorgung des 21. Jahrhunderts betrachtet [47]. Ungefähr 25 % der CHDs umfassen kritische Herzfehlbildungen, welche eine unmittelbare operative und

palliative¹ Intervention in dem ersten Lebensjahr erfordern [48, 49]. CHDs sind eine der häufigsten Ursachen von Kindersterblichkeit [50 - 52], was der Notwendigkeit einer adäquaten Früherkennung, wie bspw. durch Echokardiographie und möglicherweise zukünftig mittels genetischer Verfahren, eine wesentliche Bedeutung zukommen lässt.

1.3.1 Arten kongenitaler Herzfehlbildungen und deren

Pathophysiologie

CHDs lassen sich grundsätzlich in (1) zyanotische Herzfehlbildungen und in (2) nicht-zyanotische Herzfehlbildungen einteilen [53]. Erstere rufen eine Sauerstoffuntersättigung durch einen Rechts-Links-Shunt oder durch eine Fehlstellung der großen Arterien hervor und müssen daher unmittelbar klinisch behandelt werden [13]. Es lassen sich hunderte verschiedene anatomische CHD-Subtypen unterscheiden [53]. Der Fokus der vorliegenden Arbeit konzentrierte sich auf Patienten mit schwerwiegenden (zyanotischen) CHDs, insbesondere dem hypoplastischen Linksherzsyndrom (HLHS). Die Pathophysiologie der in dieser Arbeit analysierten CHD-Fälle wird in den folgenden Absätzen nach dem Lehrbuch von Aпитz (2013) „Pädiatrische Kardiologie: Erkrankungen des Herzens bei Neugeborenen, Säuglingen, Kindern und Heranwachsenden“ [13] beschrieben.

Nicht-zyanotische Herzfehlbildungen, die Teil der Analysen dieser Arbeit waren, umfassten die Septationsdefekte **Ventrikelseptumdefekt (VSD)** und **atrioventrikulärer Septumdefekt (AVSD)** [53]. Beim VSD ist das Ventrikelseptum nicht vollständig verschlossen und es besteht eine offene Verbindung zwischen dem linken und rechten Ventrikel. Je nach Form und Ausmaß kann dieser Defekt einen Links-Rechts-Shunt mit einem erhöhten pulmonalen Blutfluss (pulmonale Hypertonie) und eine Linksherzinsuffizienz bewirken. Kleinere VSDs liegen häufig asymptomatisch vor und es kann zu einem spontanen Verschluss kommen, ohne dass ein operativer Eingriff notwendig ist. ([13] S. 350 ff.)

AVSDs stellen Endokardkissendefekte dar, bei welchen zwischen kompletten und inkompletten AVSDs unterschieden wird. Patienten mit kompletten AVSDs, welche Teil der analysierten Patienten-Kohorte dieser Arbeit waren, zeigen einen Atriumseptumdefekt (ASD), einen

¹ Palliation: Operationsverfahren zur bestmöglichen Anpassung an gegebene physiologische Verhältnisse, die Erkrankung wird durch diesen operativen Eingriff nicht geheilt [13]

Inlet-VSD² sowie eine gemeinsame primitive AV-Kanalklappe als Öffnung zwischen Atrien und Ventrikel. ([13] S. 365 ff.)

Die in der vorliegenden Arbeit analysierten Patienten mit komplettem AVSD wiesen den genetischen Befund einer Trisomie 21³ auf. Ungefähr 43 % aller Patienten mit AVSD weisen eine Trisomie 21 auf [54]. Dieses Syndrom ist durch eine Kombination von phänotypischen Eigenschaften, u.a. einer geistigen Unterentwicklung und charakteristischen Fazies, gekennzeichnet [55, 56].

Die **Trikuspidalatresie (TA)** und **Pulmonalatresie (PA_t)** wurden in dieser Arbeit unter dem Begriff „Rechtsherzfehlbildungen“ zusammengefasst. Sie zählen zu den zyanotischen Herzfehlbildungen [53].

Die TA ist durch eine fehlende Anlage (Atresie) der Trikuspidalklappe gekennzeichnet. Anstelle der Klappe befindet sich eine undurchlässige Membran zwischen rechtem Atrium und rechtem Ventrikel. Der rechte Ventrikel ist meist stark unterentwickelt und es besteht häufig ein VSD. ([13] S. 444 ff.)

Die PA_t beschreibt einen Verschluss bzw. eine fehlende Anlage der Ausflussbahn von rechtem Ventrikel zur Pulmonalarterie. Bei der PA_t wird unterschieden, ob ein VSD vorliegt oder nicht, da dies zu unterschiedlichen hämodynamischen Auswirkungen führt. ([13] S. 404 ff.)

Das **hypoplastische Linksherzsyndrom (HLHS)** gehört zu den komplexen, zyanotischen Herzfehlbildungen [53]. Es umfasst eine starke Unterentwicklung des linken Ventrikels und der Aorta *ascendens*, sowie eine Stenose oder Atresie von Aorten- und Mitralklappe [57]. Eine detaillierte Beschreibung des HLHS, welches zentrale Thematik der vorliegenden Arbeit ist, erfolgt in Abschnitt 1.5.

In dieser Arbeit lag bei einem Patienten mit HLHS zusätzlich eine **Transposition der großen Arterien (TGA)** vor, welche zu den komplexen zyanotischen Herzfehlbildungen zählt [53].

Normalerweise sind die Pulmonalarterie mit dem rechten Ventrikel und die Aorta mit dem linken Ventrikel verbunden, bei der TGA sind die großen Arterien umgekehrt positioniert. Die häufigste Form stellt der Dextro-Typ (d-TGA) dar, bei welchem der Ursprung der Aorta anterior und rechts-seitig zum Ursprung der Pulmonalarterie liegt. Bei der TGA sind Lungen- und Körperkreislauf vollständig voneinander getrennt, sodass postnatal unmittelbar eine medikamentöse oder interventionelle Behandlung erforderlich ist. ([13] S. 418 ff.)

² Die Öffnung beim Inlet-VSD besteht im Bereich des Einlasseptums der rechten Herzkammer [13].

³ Bei der Trisomie 21 sind das gesamte Chromosom 21 bzw. Teile davon dreifach vorhanden [557].

Weitere angeborene Vitien, die Bestandteile der Analysen dieser Arbeit waren, umfassten das kongenitale **linksventrikuläre Aneurysma (LV-An)** und die familiäre **dilatative Kardiomyopathie (DCM)**. Das LV-An stellt eine idiopathische Dysplasie des linksventrikulären Endokards und Myokards dar [58]. Es liegt eine pathologische, nicht-kontraktile Ausbuchtung der ventrikulären Herzwand vor [58]. Eine vermehrte Füllung der Ausbuchtung mit Blut kann zu einer Blockade des Ausflusses aus dem linken Ventrikel führen, sodass der Körperkreislauf nicht aufrechterhalten werden kann [58].

Die DCM ist eine krankhafte Erweiterung des Myokards, insbesondere des linken Ventrikels. Die Folge ist eine ineffiziente Pumpfunktion des Ventrikels mit konsekutiv verringerter Auswurfleistung, sodass eine Unterversorgung der Organe besteht. ([13] S. 418 ff.)

1.3.2 Epidemiologie kongenitaler Herzfehlbildungen

Mit einer Prävalenz von 5 - 10 neuen Fällen pro 1000 Lebendgeburten mit moderaten bis schwerwiegenden CHDs [59 - 62] stellen CHDs eine der häufigsten Formen angeborener Fehlbildungen dar [47]. Zählt man das Auftreten der bikuspiden Aortenklappe (BAV) hinzu, welche potenziell operationsbedürftige Auswirkungen haben kann, jedoch häufig undiagnostiziert bleibt, so erhöht sich die Prävalenz auf ca. 19/1000 Lebendgeburten [63]. Abhängig davon, welche Arten von Herzfehlbildungen in Betracht gezogen werden – wie (silenter) persistierender Ductus *arteriosus* (PDA) oder kleine, muskuläre VSDs – kann insgesamt von einer Prävalenz dieser Fehlbildungen von 5 bis 50 Fällen pro 1000 Lebendgeburten ausgegangen werden [63, 64]. Die Prävalenz der CHDs erhöht sich zusätzlich, wenn ungeborene Föten mit diagnostizierten CHDs einbezogen werden [65]. Zu einer Vielzahl milder Formen von CHDs liegen jedoch keine adäquaten *Follow-up*-Studien vor und einige CHDs heilen spontan aus, sodass die reale Prävalenz ungeklärt bleibt [64].

Betrachtet man das erhobene Wiederholungsrisiko für CHDs verschiedener Studien [66 - 69], so kann ein Anstieg des Wiederholungsrisikos für jedes weitere betroffene Geschwisterkind beobachtet werden, mit bis zu 10 % bei VSD und HLHS. Ein 12-fach erhöhtes Wiederholungsrisiko kann bei Zwillingen desselben Geschlechts im Vergleich zu Zwillingen unterschiedlichem Geschlechts festgestellt werden [62, 70]. Tendenziell ist ein höheres Wiederholungsrisiko bei HLHS, Aortenklappenstenose und Koarktation der Aorta (CoA, Einengung der Aorta am Aortenbogen) im Vergleich zu anderen CHD-Typen zu beobachten [71]. Insgesamt liegt das Wiederholungsrisiko für CHDs mit 2 - 10 % [66, 67, 72]

jedoch wesentlich niedriger, als bei Mendelschem Vererbungsmodell (25 % bei autosomal rezessivem Erbgang) zu erwarten wäre. Mögliche Vererbungsmechanismen und bisherige Erkenntnisse zu Einflussfaktoren in der Ätiologie von CHDs werden daher im folgenden Abschnitt näher beschrieben.

1.4 Ätiologie kongenitaler Herzfehlbildungen

In Anbetracht des räumlich und zeitlich fein regulierten genetischen und epigenetischen Programms der Herzentwicklung [36 - 39] können sowohl genetische als auch epigenetische Einflussfaktoren in der Entstehung von CHDs von Bedeutung sein. Solche epigenetischen Faktoren beschreiben zelluläre Mechanismen mit einem Einfluss auf die Genexpression ohne eine Veränderung der DNA-Sequenz [73]. Im Hinblick auf das niedrige Wiederholungsrisiko von 2 - 10 % [66, 67, 72] wurden verschiedene Theorien zu Vererbungsmechanismen von CHDs diskutiert. Unter anderem wurden ein maßgeblicher Einfluss von *de novo* Mutationen (neu entstandene Mutationen, ohne Vererbung durch ein Elternteil) [74], eine reduzierte Penetranz und ein polygenes und/oder multifaktorielles Vererbungsmodell [69, 75 - 77] als mögliche Mechanismen vermutet. Die reduzierte Penetranz beschreibt das Vorliegen eines krankheits-assoziierten Genotyps bei Individuen, die keine Manifestation der Erkrankung zeigen. Beispielsweise treten solche Mutationen nach autosomal dominantem Vererbungsmodell mit reduzierter Penetranz in der ersten und dritten Generation bei Fällen mit CHDs auf, während in der zweiten Generation dieser Familie die vererbte Mutation keine phänotypische Auswirkung zeigt [78]. Im Folgenden werden bisher bekannte genetische und epigenetische Einflüsse in der Ätiologie von CHDs beschrieben.

1.4.1 Genetische Faktoren bei kongenitalen Herzfehlbildungen

In den letzten Jahren wurden die klassischen genetischen Ansätze, wie *Linkage*-Analysen und Resequenzierungen von Kandidatengenen, durch genomweite Hochdurchsatz-Technologien, d.h. Array-basierte Techniken und *Next Generation Sequencing* (NGS)-Verfahren, erweitert [47]. So ließen sich bisher für 15 – 20 % der CHD-Fälle eine genetische Ursache oder teratogene Einflüsse feststellen [47, 66, 69, 77]. Einen Großteil dieser CHD-Fälle mit bekannter genetischer Ursache stellen syndrom-assoziierte CHDs, d.h. CHDs mit zusätzlichen extrakardialen Manifestationen, dar [79]. Die Ätiologie nicht-syndromaler CHDs ist dagegen meist

unbekannt [69, 77]. Ungefähr 8 – 12 % der CHD-Fälle liegen in Assoziation mit großen chromosomalen Anomalien und Aneuploidien vor [66, 80, 81]. Weitere 3 – 5 % der CHD-Fälle lassen sich durch einzelne Genvarianten und Mikrodeletionen im Zusammenhang mit Mendelschen Syndromen erklären [64, 68]. Im Folgenden werden auf chromosomale Anomalien und Keimbahnmutationen in Assoziation mit CHDs, als auch auf den Einfluss möglicher somatischer, d.h. herzwegebespezifischer, Mutationen eingegangen.

1.4.1.1 Chromosomale Anomalien und CHDs

Chromosomale Aneuploidien, als erste identifizierte genetische Ursache von CHDs, stellen einen der Hauptmechanismen der Pathogenese von CHDs dar [47, 77]. Viele **chromosomale Syndrome** treten in Assoziation mit CHDs auf. Es lassen sich u.a. bei 40 – 50 % der Fälle mit Trisomie 21 [82], bei 20 – 50 % der Fälle mit Turner Syndrom [83] und bei nahezu allen Fällen mit Trisomie 13 und Trisomie 18 [84] CHDs diagnostizieren. Solche chromosomale Anomalien zeigen eine Assoziation zu verschiedenen CHDs mit unterschiedlicher Häufigkeit, sodass einer bestimmten Chromosomenaberration kein spezifischer CHD-Phänotyp zugeordnet werden kann [85]. Insgesamt liegen bei 8 - 10 % der CHD-Fälle eine Assoziation mit großen (≥ 100 kbp) chromosomalen Deletionen, Duplikationen oder auch Aneuploidien vor [66, 80]. Verschiedenste Syndrome, die durch Kopienzahlvariationen (CNVs) ausgelöst werden, treten in Zusammenhang mit CHDs auf. Die häufigste CHD-assoziierte Kopienzahlvariation (CNV) stellt eine 3-Mb-Deletion auf Chromosom 22q11.2 (Mikrodeletionssyndrom 22q11.2) dar. Diese CNV liegt bei 15 % der Fälle mit Fallot'scher Tetralogie⁴ (TOF) vor [86].

Derzeitige Analysen konnten multiple CNVs identifizieren, welche in ursächlichem Zusammenhang mit isolierten, d.h. **nicht-syndromalen CHDs**, stehen. Beispielsweise ließen sich große *de novo* CNVs bei isoliert auftretenden Fällen mit TOF [87, 88], bei linksseitigen CHDs (LS-CHDs) wie dem HLHS [89, 90], oder anderen sporadisch auftretenden CHD-Fällen detektieren [91, 92, 93].

⁴ Kongenitale Herzfehlbildung bestehend aus vier Herzdefekten (daher Tetralogie): Pulmonalstenose, Malalignment-VSD, überreitende Aorta und konsekutive Rechtsherzhypertrophie (pathologische Vergrößerung (Hypertrophie) des Muskelgewebes im rechten Ventrikel) [13].

1.4.1.2 Keimbahnmutationen und CHDs

Verschiedenste Einzelgen-Mutationssyndrome treten in Zusammenhang mit CHDs (im Folgenden als **syndrom-assoziierte CHDs** bezeichnet) auf. Beispielsweise konnten *Gain-Of-Function*-(GOF)-Mutationen in *PTPN11* [94], *SOS1* [95] und *RAF1* [96] in ursächlichem Zusammenhang mit dem Noonan-Syndrom, welches bei 80 – 90 % der Patienten [97, 98] in Assoziation mit CHDs auftritt, identifiziert werden. Das Alagille-Syndrom, welches bei 90 % der Patienten in Assoziation mit CHDs auftritt [99 - 101], wird hauptsächlich durch dominante *Loss-Of-Function*-(LOF)-Mutationen in *JAG1* ausgelöst [102]. Das Catel-Manzke-Syndrom liegt bei 40 % der Fälle in Assoziation mit einer Herzfehlbildung (VSD, ASD) vor [103]. Bei diesem Syndrom handelt es sich um eine seltene Knochenkrankheit, die durch eine beidseitige Hyperphalangie und Klinodaktylie des Zeigefingers charakterisiert ist, sowie, in typischen Fällen, in Verbindung mit einer Pierre-Robin-Sequenz (Gaumenspalte, Mikrogenie und Glossoptose) auftritt [103]. In der vorliegenden Arbeit konnte im Rahmen der Studie von Ehmke et al. (2014) [104] das *TGDS*-Gen als kausales Gen für das beobachtete autosomal-rezessive Vererbungsmodell des Catel-Manzke-Syndroms identifiziert werden.

In ursächlichem Zusammenhang mit **nicht-syndromalen CHDs** konnten verschiedenste Gene, die u.a. in kardialen Signalwegen und der transkriptionellen Regulation der Herzentwicklung involviert sind, identifiziert werden [79]. Unter anderem ließen sich Mutationen in TF-kodierenden Genen, wie *NKX2.5* [105 - 108] und *GATA4* [109 - 111] in Assoziation mit nicht-syndromalen CHDs nachweisen. Die Haploinsuffizienz⁵ kardialer TFs könnte hierbei möglicherweise in der Pathogenese von CHDs relevant sein [112 - 114].

Kürzlich konnte in einer WES-Studie von 1823 CHD-Familien (Sifrim und Hitz et al., 2016⁶), an der sich im Rahmen dieser Arbeit beteiligt wurde, eine signifikante Anreicherung von *de novo* LOF-Mutationen in bekannten CHD-Genen bei syndrom-assoziierten CHDs nachgewiesen werden. Hingegen waren bei nicht-syndromalen CHDs vererbte LOF-Mutationen in CHD-Genen, die gleichsam in den phänotypisch gesunden Eltern vorlagen, signifikant angereichert (Sifrim und Hitz et al., 2016⁶). Somit zeigen syndrom-assoziierte und nicht-syndromale CHDs ein distinktes genetisches Modell, mit einer entscheidenden Bedeutung von *de novo* Mutationen bei syndrom-assoziierten CHDs und vererbten proteintrunkierenden Genvarianten bei nicht-syndromalen CHDs (Sifrim und Hitz et al., 2016⁶).

⁵ Der diploide Organismus hat hierbei nur eine funktionelle Kopie eines Gens (das andere ist durch eine Mutation inaktiviert), sodass im haploiden Zustand der normale Phänotyp nicht aufrechterhalten werden kann [557].

⁶ Dokument derzeit in Revision, eingereicht bei Nature Genetics (siehe Abschnitt 11)

1.4.1.3 Somatische Mutationen und CHDs

Somatische Mutationen, d.h. im Herzgewebe selbst, könnten einen alternativen genetischen Mechanismus in der Entstehung von CHDs darstellen [78]. In *GATA6* konnten bspw. kürzlich zwei bisher unbekannte heterozygote Mutationen im Herzgewebe von Patienten mit TOF nachgewiesen werden, die weder in peripherem Blut dieser Patienten, noch in Herzgewebeproben einer Kontrollkohorte nachweisbar waren [115]. Andere (nicht-kongenitale) Herzerkrankungen zeigten ebenfalls somatische Mutationen. In Genen kodierend für Connexin-40 (*GJA5*) und Cx43 (*GJA1*) ließen sich herzspezifische, somatische Mutationen in sporadischen Fällen mit Vorhofflimmern detektieren [116, 117]. Insgesamt ist jedoch die bisherige Datenlage zur Erforschung somatischer Mutationen in Zusammenhang mit kardialen Defekten nur gering ausgeprägt, zumal es sich beim Herzgewebe um ein schwer zugängliches Gewebematerial handelt.

Nach Schätzungen durch gegenwärtige Sequenzierungs-Strategien mittels NGS-Technologien treten in jeder Generation ca. ≤ 10 seltene, Proteinfunktion-beeinträchtigende *de novo* Mutationen auf [118, 119]. Zwar basieren diese Daten auf Analysen von Blutproben, die eine deutlich höhere Zellteilungsfrequenz im Vergleich zu differenziertem Herzgewebe und damit ein erhöhtes Risiko für das Auftreten von *de novo* Mutationen haben, jedoch könnten solche somatischen Mutationen in frühen proliferativen Phasen der kardialen Entwicklung auftreten und so in der Entstehung von CHDs eine Rolle spielen. Obwohl die Daten von Choi et al. (2009a) [118] und Sanders et al. (2012) [119] keine direkte Schätzung von Mutationsfrequenzen in proliferierenden und differenzierten Zellen zulassen, so verdeutlichen sie dennoch, dass *de novo* Mutationen frequentiert auftreten. Somit können Analysen von somatischen *de novo* Mutationen im Herzgewebe von CHD-Patienten potenziell informativ sein. Um einen solchen alternativen genetischen Mechanismus in der Entstehung des HLHS zu untersuchen, wurden in der vorliegenden Arbeit Herzgewebeproben von HLHS-Patienten analysiert.

1.4.2 Epigenetische Aspekte kongenitaler Herzfehlbildungen

Der Begriff „Epigenetik“, erstmalig durch Conrad Waddington (1942) beschrieben [120], lässt sich nach heutiger Definition zusammenfassen als modifizierbare, zelluläre Mechanismen mit einem Einfluss auf die Genexpression, die ohne eine Veränderung der DNA-Sequenz einhergehen und mitotisch sowie meiotisch vererbbar sein können [73]. Epigenetische

Regulationsmechanismen umfassen Modifikationen von DNA-Basen (wie die DNA-Methylierung), post-translationale Modifikationen von Histonproteinen, RNA-basierte Mechanismen im Nukleus und nukleäre Positionseffekte [121]. Dieses genomweite epigenetische Muster an Modifikationen, das sogenannte Epigenom, unterliegt präzisen koordinierten Veränderungen in definierten Entwicklungsstadien und ist in der Regulation von multiplen physiologischen Prozessen während der menschlichen Entwicklung involviert [122]. Epigenetische Veränderungen tragen zu zell- und gewebespezifischer Genexpression bei [123 - 125]. Umweltfaktoren können diese epigenetischen Modifikationen beeinflussen [126, 127]. Der maßgebliche Einfluss von **Umwelt- und *in utero*-Faktoren** liegt wahrscheinlich innerhalb der perikonzeptionellen Periode (drei Monate vor der Schwangerschaft bis zum Ende des ersten Trimenons) mit der kritischen Phase der kardialen Entwicklung zwischen der 2. und 7. Gestationswoche [59, 128].

Beispielsweise konnte ein bis zu fünffach erhöhtes Risiko für CHDs bei prägestationalem maternalen Diabetes mellitus beobachtet werden [129 - 131]. Epigenetische Mechanismen, wie die DNA-Methylierung, ließen sich in verschiedenen Studien als Einflussfaktoren in der Entstehung von Typ 2 Diabetes identifizieren [132 - 134]. Weitere Risikofaktoren für die Entstehung von CHDs umfassen Umwelt-Teratogene, wie Dioxine, polychlorierte Biphenyle und Pestizide [135]. Beispielsweise können Histon-Deacetylasen, die in der Histonmodifikation und somit in der Regulation der Transkription involviert sind, durch teratogene Ereignisse inhibiert werden. Auch parentaler Alkohol-, Kokain- und Tabakkonsum stellen bekannte Risikofaktoren für CHDs in der Nachkommenschaft dar [136 - 139]. Eine parentale Exposition zu Alkohol steht in ursächlichem Zusammenhang mit fetalen Alkohol-Spektrum-Störungen (FASD), die verschiedene körperliche und kognitive Entwicklungsstörungen umfassen, wie auch CHDs. Beispielsweise lassen sich Veränderungen des DNA-Methylierungsprofils in frühen Entwicklungsstadien bei Alkoholexposition in Mausembryos detektieren [140].

Ein Risikofaktor für beide Geschlechter stellt das parentale Alter dar, indem bei älteren Müttern häufig eine TGA und bei älteren Vätern eine höhere Inzidenz von VSDs in der Nachkommenschaft identifiziert werden konnten [141 - 144]. Veränderungen des DNA-Methylierungsmusters mit zunehmendem Alter eines Individuums ließen sich bisher in verschiedenen Studien nachweisen [145 - 152].

1.4.2.1 DNA-Methylierung als epigenetischer Regulationsmechanismus

In der vorliegenden Arbeit wurden die Aspekte der DNA-Methylierung im Herzgewebe von Patienten mit unterschiedlichen CHDs, insbesondere dem HLHS, analysiert. Daher wird im Folgenden auf die DNA-Methylierung als epigenetischer Regulationsmechanismus, sowie auf die bisherigen Erkenntnisse zu molekularen Mechanismen der DNA-Methylierung in der Entstehung von CHDs eingegangen.

Die DNA-Methylierung beschreibt eine kovalente DNA-Modifikation durch Methylierung der DNA, d.h. primär die Methylierung eines Cytosins im Kontext eines Cytosin-Guanin-Dinukleotids (CpG) [153]. Das methylierte Cytosin (5-mC) wird auch als „fünfte Base“ der DNA bezeichnet [154]. Ungefähr 70 – 80 % aller menschlichen CpG-Loci liegen methyliert vor [155]. Neben der Methylierung im CpG-Kontext können auch Cytosine gefolgt von anderen Basen, wie z.B. einer Thymin- oder Adenin-Base, eine Methylierung aufweisen [156]. Dies wird als nicht-CpG-Methylierung bezeichnet und kann bspw. in kultivierten embryonalen Stammzellen detektiert werden [156]. DNA-Methyltransferasen (DNMTs) katalysieren den Transfer einer Methylgruppe von S-Adenosyl-Methionin (SAM) auf das Cytosin [157]. Die DNA-Methyltransferase 1 (DNMT1) ist hierbei primär in dem Erhalt der DNA-Methylierung während der Zellteilung involviert [158]. Für die *de novo* DNA-Methylierung, welche bspw. in frühen Entwicklungsstadien von funktioneller Bedeutung ist, sind die DNA-Methyltransferasen DNMT3A und DNMT3B essenziell [159]. Die DNA-Methylierung stellt eine reversible Modifikation dar, die passiv oder aktiv entfernt werden kann [160].

Innerhalb des Genoms sind CpGs nicht gleichmäßig verteilt, sondern treten gehäuft in sogenannten CpG-Inseln (CGIs) auf [161]. CGIs sind DNA-Regionen von > 200 Basenpaaren (bp) mit > 50 %igem GC-Gehalt und einem Verhältnis von erwartetem zu tatsächlichem CpG-Gehalt von 0,6 [161]. CGIs zeigen eine Kollokalisierung zu Promotoren konstitutionell exprimierter Gene und ca. 40 % dieser Gene weisen ein gewebespezifisches Expressionsprofil auf [162, 163]. Die DNA-Methylierung von Promotorregionen zeigt häufig eine negative Korrelation mit der Expression von *downstream* gelegenen Genen [164 – 166]. Eine Inhibition der Genexpression kann hierbei z.B. durch die sterische Hinderung der Bindung von Transkriptionsfaktoren (TFs) an Transkriptionsfaktorbindestellen (TFBSs) in der Promotorregion bewirkt werden, wobei die Transkriptionshemmung hauptsächlich über die Rekrutierung von *Methyl-CpG Binding Domain* (MBD)-Proteinen erfolgt [167]. Die Affinität von TFs zu TFBSs kann somit durch die DNA-Methylierung beeinflusst werden [168] und die Bindung von TFs an TFBSs verhindern [168 - 171]. Eine differenzielle DNA-Methylierung zwischen Zellen oder Geweben

innerhalb eines Individuums, oder zwischen unterschiedlichen Individuen lässt sich innerhalb sogenannter differenziell methylierter Regionen (DMRs) detektieren [172]. Diese umfassen DNA-Bereiche, die bei multiplen benachbarten CpGs eine differenzielle Methylierung aufweisen. Bisher ließen sich verschiedene DMRs nachweisen, wie u.a. auch reprogrammierungsspezifische DMRs (RDMRs), welche im Rahmen der somatischen Zellreprogrammierung eine differenzielle Methylierung zeigen und die Unterscheidung zwischen verschiedenen Gewebetypen ermöglichen [173].

Die DNA-Methylierung ist in einer Vielzahl von biologischen Prozessen involviert, wie der X-Chromosom-Inaktivierung, dem genomischen Imprinting und auch der Embryonalentwicklung [174]. In der **Kardiogenese** hat die DNA-Methylierung eine wesentliche Funktion in der Entwicklung und Reifung von Kardiomyozyten [38]. Dynamische Veränderungen des Methylierungsmusters lassen sich in Zusammenhang mit einer alterierenden Transkription von Genen der Kardiogenese detektieren [39]. In der Studie von Chamberlain et al. (2014) [39] ließ sich eine herunterregulierte Genexpression des *HAS2*-Gens im Herzen von Mausembryonen in Zusammenhang mit einer erhöhten DNA-Methylierung des assoziierten Enhancers nachweisen. Solche Enhancer-Regionen regulieren die Genexpression durch räumliche Orientierung aktivierender Faktoren an Promotorregionen über chromosomales *Looping* [175]. Die DNA-Methylierung in Enhancer-Regionen ist mit Veränderungen der Genexpression assoziiert [176]. Insgesamt stellen epigenetische Modifikationen, wie die DNA-Methylierung, einen wichtigen Mechanismus in der Regulation der Herzentwicklung dar [36, 37].

Bisher liegen nur wenige Studien zur **DNA-Methylierung im Herzgewebe von Patienten mit Herzfehlbildungen** vor. Ein aberrantes DNA-Methylierungsmuster in linksventrikulärem Gewebe konnte bei Patienten mit terminaler Herzinsuffizienz [177] sowie bei Fällen mit DCM [178] in Korrelation mit veränderter Genexpression detektiert werden. Somit können Erkrankungen des Myokards mit Veränderungen der DNA-Methylierung einhergehen. Bei Patienten mit TOF konnte in der Studie von Sheng et al. (2013) [179] eine aberrante DNA-Methylierung von *NKX2.5* und *HAND1* mit einer negativ korrelierten mRNA⁷-Expression dieser Gene in kardialem Gewebe (rechter Ventrikel) detektiert werden, sodass Veränderungen der DNA-Methylierung zu einer reduzierten Transkription essenzieller Herzentwicklungs-Gene beitragen können. Des Weiteren konnte in einer Fall-Kontroll-Studie von Obermann-Borst et al. (2011) [180] ein hohes Level von Biomarkern wie SAM (Methylgruppendonor für

⁷ messenger RNA, einzelsträngiges RNA-Transkript eines Gens, das im Rahmen der Translation zur Synthese von Proteinen dient [557]

DNMTs) und Folat (als Vorstufe von Tetrahydrofolat, einem Lieferanten von Methylgruppen) in Assoziation mit komplexen syndromalen CHDs nachgewiesen werden.

1.5 Das hypoplastische Linksherzsyndrom

Das hypoplastische Linksherzsyndrom (HLHS), ein kongenitales Herz- und Gefäßfehlbildungssyndrom [57, 181], stellt einen der schwerwiegendsten CHDs mit den höchsten Mortalitätsraten dar [182 – 184]. Insgesamt lassen sich 25 - 40 % aller neonatalen kardialen Todesfälle auf das HLHS zurückführen [183, 184]. Das HLHS ist die häufigste Form univentrikulärer Herzfehlbildungen [185], d.h. Herzfehler, bei denen nur ein Ventrikel funktionsfähig ist (der andere Ventrikel ist unterentwickelt, sodass dieser seine Funktion nicht aufrechterhalten kann) [13]. Es tritt mit einer Prävalenz von 2 - 3 Fällen pro 10.000 Lebendgeburten auf [183, 186, 187]. Zieht man spontane und induzierte Aborte betroffener Feten in Betracht, so erhöht sich die Prävalenz [182, 187]. Zur Entstehung des HLHS ist bis heute nur wenig bekannt [188].

1.5.1 Anatomie und Pathophysiologie des HLHS

Das HLHS beschreibt anatomisch ein Spektrum kardialer Fehlbildungen der linken Herzhälfte und angrenzender Gefäßstrukturen, die eine signifikante Hypoplasie (Unterentwicklung) des linken Ventrikels, eine Atresie (Verschluss), Stenose (Verengung) oder Hypoplasie der Aorten- und/oder Mitralklappe, sowie eine Hypoplasie der Aorta *ascendens* und des Aortenbogens umfassen [57, 181]. Entsprechend der pathogenetischen Klassifizierung nach Clark (1996) [189] lässt sich das HLHS den „Anomalien des intrakardialen Blutflusses“ zuordnen und stellt nach EPCC-Klassifikationssystem (European Paediatric Cardiac Code) [53] einen Subtyp linksventrikulärer Ausflusstrakt-Obstruktionen (LVOTOs) dar.

Bisher bekannte anatomische und pathophysiologische Gegebenheiten des HLHS werden in den folgenden Absätzen anhand des Lehrbuchs von Aпитz (2013) „Pädiatrische Kardiologie: Erkrankungen des Herzens bei Neugeborenen, Säuglingen, Kindern und Heranwachsenden“ [13] beschrieben.

Das Ausmaß der Hypoplasie von linksventrikulären und aortalen Strukturen beim HLHS reicht von einem geringfügig vorhandenen Blutfluss durch den linken Ventrikel bei MS/AS bis zu

einem, im Extremfall nur reiskorngroßen, linken Ventrikel ohne Blutfluss bei MA/AA. Die Aorta *ascendens* ist daher unterentwickelt. Bei Durchgängigkeit der Mitralklappe liegt teilweise eine Endokardfibroelastose (EFE) des linken Ventrikels vor. ([13] S. 374 ff.)

Klinisch werden folgende anatomische Varianten beim klassischen HLHS entsprechend der Ausprägung der Herzklappen der linken Herzhälfte klassifiziert [190, 191] (Abbildung 3):

- a) Mitralatresie/Aortenatresie (MA/AA)
- b) Mitralstenose/Aortenatresie (MS/AA)
- c) Mitralstenose/Aortenstenose (MS/AS)

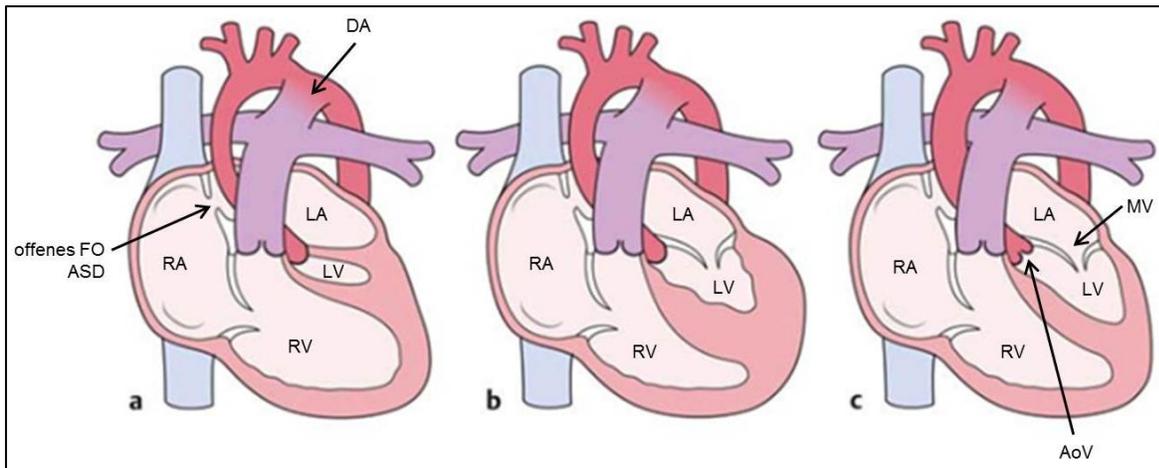


Abbildung 3 Anatomische Varianten des HLHS. a: Mitral- und Aortenatresie, b: Mitralstenose und Aortenatresie, c: Mitral- und Aortenstenose. FO: Foramen ovale, DA: Ductus arteriosus, AoV: Aortenklappe, MV: Mitralklappe, LA: linkes Atrium, LV: linker Ventrikel (Abbildung aus Haas and Kleideiter (2011) [192], modifiziert)

Die linke Herzhälfte ist beim HLHS aufgrund der starken Unterentwicklung seiner Strukturen, d.h. Mitralklappe, Aortenklappe, *Cavum* und linksventrikulärer Ausflusstrakt (LVOT), nicht in der Lage, ein ausreichendes Herzzeitvolumen (HZV) zu fördern. Somit besteht eine Unterversorgung des Körperkreislaufs. Das Überleben nach der Geburt hängt von zwei anatomischen Strukturen ab, die intrauterin bei allen Menschen bestehen: dem DA und dem FO (Abbildung 3 und siehe Abschnitt 1.1). Über den offenen DA erfolgt eine systemische Perfusion, wobei für das postnatale Überleben ein Offenbleiben des DAs und ein ausreichender Links-Rechts-Shunt auf Vorhofebene erforderlich sind. Beim HLHS kann sich das oxygenierte Blut aus den Lungenvenen nicht über den linken Ventrikel in den Körperkreislauf entleeren, sondern fließt über ein offenes FO bzw. über einen ASD aus dem linken Atrium in das rechte Atrium (Links-Rechts-Shunt) ab. Im rechten Atrium mischt sich das oxygenierte Blut mit dem desoxygenierten Blut der Hohlvenen. Dieses Mischblut gelangt dann über den rechten Ventrikel,

der somit eine erhöhte Volumenbelastung aufweist, in die Pulmonalarterie und von dort teils in den Lungen- und teils in den systemischen Kreislauf über den DA. ([13] S. 374 ff.)

Bleibt das HLHS unbehandelt, so führt die Lungenüberflutung durch den postnatal verringerten pulmonalvaskulären Widerstand zu einer Abnahme des Rechts-Links-Shunts über den DA, wodurch eine Verringerung des systemischen HZVs hervorgerufen wird. Darüber hinaus bewirkt die physiologische postnatale Konstriktion des DAs eine drastische Reduktion der Systemperfusion und der Perfusion der Koronararterien. Über eine Steigerung des Sympathikotonus führt dies zum Anstieg des systemischen Gefäßwiderstands und somit zum Erhalt eines adäquaten Blutdrucks. Die Folgen dieser Mechanismen sind ein kardialer Schock und die Ausbildung einer schweren metabolischen Azidose und atrialer Hypoxämie. ([13] S. 374 ff.)

Die **Therapie** umfasst unmittelbar nach Diagnosestellung ein Offenhalten des DAs mittels Prostaglandin-E-Infusion, um die ductusabhängige Körper- und Koronarperfusion zu gewährleisten. Des Weiteren muss eine ausreichend große interatriale Kommunikation vorliegen, um einen freien Abfluss des oxygenierten Blutes aus den Lungenvenen zu gewährleisten. Die operative Standardtherapie ist beim HLHS eine dreistufige Palliation, dessen Resultat eine komplette Kreislauftrennung von Lungen- und Körperkreislauf nach dem sogenannten Fontan-Prinzip ist. Bei der resultierenden „Fontan-Zirkulation“ fließt das gesamte venöse, desoxygenierte Blut der Hohlvenen – ohne zuvor einen Ventrikel zu passieren – direkt in die Lungenstrombahn, sodass eine passive Lungenzirkulation besteht. Nach Anreicherung mit Sauerstoff gelangt das Blut über die Pulmonalvenen und den ASD in den rechten Ventrikel, welcher nun mit der Aorta verbunden ist und somit den Körperkreislauf versorgt (Details siehe Abschnitt 8 Tabelle 30). Des Weiteren kann eine Herztransplantation als therapeutische Maßnahme durchgeführt werden, wobei neben dem Herzen auch die *Aorta ascendens* und ein Teil des Aortenbogens ersetzt werden müssen. Jedoch ist diese Therapiemöglichkeit aufgrund des Mangels von Spenderorganen begrenzt. ([13] S. 379 ff.)

In den letzten 30 Jahren hat sich das klinische Management aufgrund technischer und strategischer Fortschritte drastisch geändert, sodass beim HLHS Überlebensraten bis zum fünften Lebensjahr von > 70 % im Vergleich zur 100 %igen Mortalität bei unbehandelten Neonaten erreicht werden können [183, 185, 193, 194]. Jedoch birgt die dreistufige Palliation ein hohes Mortalitätsrisiko, mit mehreren Operationen am offenen Herzen und perioperativen Komplikationen [195], sodass alternative Behandlungsstrategien etabliert werden müssen. Zur Erweiterung der Therapiemöglichkeiten, insbesondere in Hinblick auf die anatomische

Heterogenität des HLHS, ist ein genaues Verständnis über die Ätiologie und die Erforschung der Pathomechanismen von wesentlicher Bedeutung.

1.5.2 Ätiologie des HLHS

Der zugrundeliegende Pathomechanismus des HLHS ist bisher weitgehend unklar [188]. Die Mehrzahl der HLHS-Fälle tritt sporadisch ohne eine bekannte Ursache auf [78, 196]. Verschiedenste Theorien wurden bezüglich der Ätiologie dieser komplexen Herzfehlbildung diskutiert.

Während der fetalen Zirkulation wird der linke Ventrikel primär durch einen Blutfluss über das FO versorgt [11]. Eine Beeinträchtigung dieses Rechts-Links-Shunts, und somit ein Ausbleiben eines adäquaten Blutflusses, könnte eine linksventrikuläre Wachstumsminde rung zur Folge haben [188]. Diese „Fluss-Theorie“ wird auch durch die Beobachtung unterstützt, dass bei Feten mit HLHS kleinere FOs vorliegen, als bei normalen gesunden Feten [197]. Zudem konnte das Auftreten anatomischer Anomalien des interatrialen Septums (IAS) in Assoziation mit dem HLHS detektiert werden [198]. Auch ein intaktes IAS ohne Öffnung (FO), dessen Offenbleiben wichtige Voraussetzung des fetalen Kreislaufs ist, konnte *in utero* in Assoziation mit dem HLHS beobachtet werden [199].

Eine prävalente Theorie zur Ätiologie des HLHS beinhaltet die Hypothese, dass “primäre“ anatomische Defekte des linksventrikulären Ein- und Ausflusstrakts einen aberranten Blutfluss durch die linke Herzhälfte bewirken, sodass sekundär konsekutiv eine Unterentwicklung linksventrikulärer und aortaler Strukturen resultiert [200]. In verschiedenen Tiermodellen konnte eine linksventrikuläre Hypoplasie als Resultat induzierter linksseitiger Obstruktionen des Ein- und Ausflusstrakts [201, 202], oder durch anatomische Modifikation des linken Vorhofs gezeigt werden [203]. Klinische Beobachtungsstudien zeigten, dass Feten mit einer in frühem Gestationsstadium entdeckte Aortenstenose pränatal eine progrediente Hypoplasie der linken Herzhälfte und neonatal einen HLHS-Phänotyp entwickelten [204]. Diese Beobachtung konnte auch im Tiermodell bestätigt werden, indem bei veränderter Blutflussdynamik eine aberrante Kardiogenese und Genexpression detektiert werden konnte [205, 206].

Evidenzbasierte Erkenntnisse lassen zunehmend eine **genetische Komponente** in der Entstehung von LS-CHDs wie dem HLHS vermuten. Ungefähr 10 % der HLHS-Fälle weisen ein **genetisches Syndrom** auf [207 - 210]. Beschrieben wurden u.a. das Turner-Syndrom [211],

Noonan-Syndrom [212], Holt-Oram-Syndrom [213], Jacobsen-Syndrom [214], Potocki-Lupski-Syndrom [215], Kabuki-Syndrom [216], Kleefstra-Syndrom [217] sowie die Trisomie 13 und 18 [208, 209] (Abbildung 4).

Die Mehrzahl **nicht-syndromaler** HLHS-Fälle tritt sporadisch auf und die **Identifizierung kausaler Gene** gestaltet sich problematisch, da häufig pro Familie nur ein Fall mit HLHS bekannt ist und die wenigen Familien mit einer erkennbaren Vererblichkeit teils eine variable Penetranz zeigten [200, 218 – 220]. Zwar konnten in einigen Fällen Genvarianten in Assoziation mit dem HLHS beschrieben werden (Abbildung 4), jedoch ist die kausative Rolle dieser Mutationen bisher noch nicht vollständig geklärt. In *NKX2.5*, kodierend für einen TF in der frühen Kardiogenese, konnte in verschiedenen Fällen mit CHDs, inklusive dem HLHS, Mutationen identifiziert werden [106, 107, 221]. In der Studie von Iascone et al. (2012) [196] ließen sich seltene Genvarianten in *FOXC2* und *FOXL1*, die in der Herzentwicklung involviert sind, in Assoziation mit dem HLHS detektieren. Auch in *GJA1* konnten Mutationen in Zusammenhang mit dem HLHS identifiziert werden [222]. In *NOTCH1*, das in der Herzklappen-Entwicklung von Bedeutung ist [223, 224], lassen sich Missense⁸- und LOF-Mutationen in LS-CHDs, inklusive dem HLHS, zeigen [196, 225, 226]. Mutagenese-Screens im Mausmodell zeigten homozygote Mutationen in *NADSYN1*, *PCDHB9* und *PEPD* in mutanten Mäusen mit einem HLHS-Phänotyp, sodass Veränderungen in Genen des Energiemetabolismus (*NADSYN1*) sowie der interzellulären und Zellmatrix-Interaktion (*PCDHB9*, *PEPD*) in der Entstehung des HLHS involviert sein könnten [227].

Linkage-Analysen zeigten eine Assoziation mit dem HLHS in den **chromosomalen Regionen** 10q22 und 6q23 und in 14q23 in Kombination mit BAV [228]. Auch die Regionen 2p15, 2p23.2, 10q21.2 und 16p12.2 konnten in Zusammenhang mit LVOTOs, inklusive dem HLHS, identifiziert werden [229] (Abbildung 4). In der Studie von Stankiewicz et al. (2009) [230] ließen sich Microdeletionen im FOX-Cluster (16q24.1) in ursächlichem Zusammenhang mit der alveolar-kapillären Dysplasie und anderen Malformationen wie dem HLHS nachweisen. Mikrodeletionen in der Region 9q34.3, welche u.a. die Gene *EHMT1* and *NOTCH1* umfasst, konnte in Assoziation mit dem HLHS – teilweise im Zusammenhang mit dem Kleefstra-Syndrom - detektiert werden [78, 87, 217]. Des Weiteren ließen sich in einer Familien-basierten CNV-Studie kosegregierende und *de novo* CNVs in Assoziation mit LS-CHDs, inklusive dem HLHS, identifizieren [89].

⁸ Missense-Mutation: sinnverändernde Mutation. Punktmutation, die den Einbau einer anderen Aminosäure in das Protein bewirkt [557].

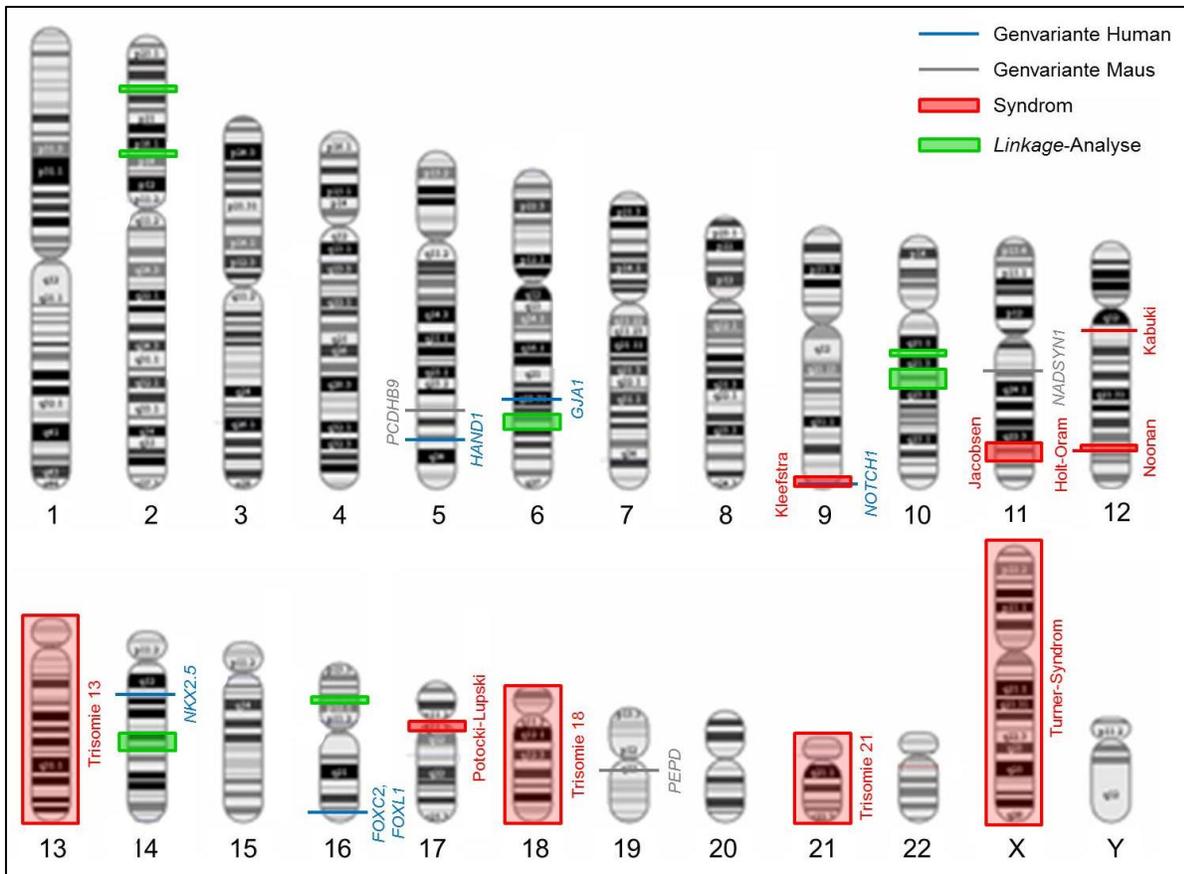


Abbildung 4 Karyogramm-Darstellung bekannter Genvarianten, chromosomalen Regionen und Syndromen, die in Assoziation mit dem HLHS bekannt sind. Aus der Literatur bekannte Gene, die bei Fällen mit HLHS im kausalen Zusammenhang mit dieser Herzfehlbildung identifiziert wurden, sind in blau dargestellt. In grau markiert sind Gene, die mit einem HLHS-Phänotyp im Mausmodell assoziiert sind. Aus *Linkage*-Analysen bekannte chromosomale Regionen, die eine Assoziation mit dem Auftreten von HLHS zeigen, sind grün hinterlegt. Chromosomale Regionen von Syndromen, die in Zusammenhang mit dem HLHS auftreten (nicht alle genannt), sind rot hinterlegt.

Grundsätzlich kann eine **familiäre Häufung** mit einer Inzidenz von CHDs von bis zu 19 % bei Familienmitgliedern von HLHS-Fällen beobachtet werden [228, 231 - 235]. Ungefähr 10 % dieser CHDs besteht aus BAV-Fällen [236 – 238], einer milden LVOTO-Ausprägung. Beim HLHS ist zudem ein im Vergleich zu anderen CHD-Typen höheres Wiederholungsrisiko [71] mit einem bis zu 10 %igem Wiederholungsrisiko für Geschwisterkinder zu beobachten [68, 234, 239, 240], sodass eine starke Evidenz für eine zugrundeliegende genetische Komponente vorliegt.

In der Studie von Kobayashi et al. (2014) [241] ließen sich **epigenetische Faktoren** in Assoziation mit dem HLHS identifizieren. Die Autoren konnten in induzierbaren pluripotenten Stammzellen (iPSCs) von HLHS-Patienten eine transkriptionelle Repression und epigenetische Modifikation von *NKX2.5*, *HAND1* und *NOTCH1*, sowie reduzierte Expressionslevel des NOTCH/HEY-Signalweges nachweisen. In der Studie von Zaidi et al. (2013) [74] ließ sich eine

signifikante Anreicherung von *de novo* Mutationen in Genen, die in der H3K4-Methylierung (bzw. Demethylierung oder Auslesung) involviert sind, in Zusammenhang CHDs, inklusive dem HLHS, nachweisen. Der epigenetische Mechanismus der Histonmodifikation von H3K4, d.h. chemische Modifikationen des Aminosäurerestes Lysin 4 vom Histonprotein⁹ H3, ist u.a. in der Regulation wichtiger Gene der Embryonalentwicklung und Kardiogenese involviert [243 - 245].

1.6 Hypothese und Zielsetzung

Während sich in den letzten Jahren wesentliche Fortschritte im klinischen Management des HLHS erzielen ließen, bleibt die Ätiologie dieser komplexen Herzfehlbildung bis heute weitgehend unklar [188]. Lediglich in wenigen Fällen konnten Genvarianten in Zusammenhang mit dem HLHS identifiziert werden [106, 107, 221, 196, 222, 225, 226], wobei die Kausalität dieser Mutationen bisher nicht eindeutig geklärt ist. In familiären Fällen mit HLHS lässt sich u.a. eine variable Penetranz detektierter Mutationen [200, 218 - 220] und eine unterschiedliche Ausprägung linksseitiger kardialer Defekte feststellen [233, 237, 246]. Dies lässt vermuten, dass zusätzlich zu einer vererbten genetischen Komponente weitere Faktoren in der Ätiologie des HLHS involviert sind. Hypothetisch könnten daher in der Entstehung des HLHS prädisponierende genetische Faktoren sowie zusätzliche, in Kombination wirkende, *de novo* Mutationen und epigenetische Faktoren eine Veränderung in der Genexpression bewirken, sodass eine pathologische Morphogenese linksseitiger Strukturen resultiert. Potenziell können sowohl Ereignisse auf Keimbahnebene als auch auf somatischer Ebene¹⁰, d.h. im betroffenen Herzgewebe, ursächlich für die Entstehung dieser komplexen Herzfehlbildung sein. Bisher liegen jedoch kaum Daten zur Erforschung von Mutationen und epigenetischen Veränderungen auf somatischer Ebene in Zusammenhang mit kardialen Defekten vor, zumal es sich beim Herzgewebe um ein schwer zugängliches Gewebematerial handelt. In der vorliegenden Arbeit erfolgten daher – neben genetischen Analysen auf Keimbahnebene – genetische, epigenetische und transkriptionelle Analysen auf somatischer Ebene. Hierfür standen Keimbahn- und Herzgewebeproben von Patienten mit CHDs, insbesondere dem HLHS, zur Verfügung.

⁹ Histone stellen Bestandteile des Chromatins dar und sind von funktioneller Bedeutung für die Verpackung der DNA sowie für die Regulation der Genexpression [242].

¹⁰ Der Begriff „somatisch“ bezieht sich in dieser Arbeit grundsätzlich auf Ereignisse im Herzgewebe und nicht in der Keimbahn.

Um die molekularen Mechanismen in der Entstehung von CHDs, insbesondere des HLHS, auf Keimbahn- und somatischer Ebene zu untersuchen, wurden in dieser Arbeit folgende Fragestellungen bearbeitet:

1. Lassen sich auf **Keimbahnebene** bei Fällen mit einem syndrom-assoziierten CHD und nicht-syndromalen CHDs (HLHS) in familiärem Kontext krankheitsrelevante oder kausale Genvarianten¹¹ identifizieren?
2. Können auf **somatischer Ebene**, d.h. im betroffenen Herzgewebe, bei Patienten mit HLHS krankheitsrelevante oder kausale Genvarianten identifiziert werden?
 - a. Zeigen die detektierten Genvarianten eine Überlappung zu bereits in der Literatur bekannten Genen, die in Assoziation mit Herzfehlbildungen identifiziert wurden?
 - b. Sind Mutationen nachweisbar, die ausschließlich im Herzgewebe, nicht jedoch in der Keimbahn der Patienten auftreten?

Treten Veränderungen in der DNA-Methylierung im Herzgewebe von Patienten mit CHDs, speziell des HLHS, in einem möglichen Zusammenhang mit der Herzfehlbildung auf?

- a. Kann eine differenzielle DNA-Methylierung in Herzgewebeproben von HLHS-Patienten im Vergleich zu Rechtsherzfehlbildungen als Kontrollkohorte nachgewiesen werden?
- b. Lassen sich verschiedene Herzregionen durch ein differenzielles DNA-Methylierungsprofil unterscheiden? Welche Gene bzw. *Gen-Pathways* zeigen eine differenzielle DNA-Methylierung?
- c. Welche Gene liegen differenziell methyliert im Herzgewebe von HLHS-Patienten im Vergleich zu gesundem fetalem Herzkontrollgewebe vor?

Zeigen Herzgewebeproben von Patienten mit HLHS eine differenzielle Genexpression im Vergleich zu Herzgewebeproben von Patienten mit Rechtsherzfehlbildungen?

3. Lassen sich in einer **integrativen Analyse** überlappende Gene mit Mutationen, differenzieller DNA-Methylierung und/oder differenzieller Genexpression in Zusammenhang mit dem HLHS identifizieren? Können hierbei Überlappungen betroffener *Gen-Pathways* auf Genom-, Epigenom- und Transkriptomebene detektiert werden?

¹¹ Definition zu „krankheitsrelevant“ bzw. „kausal“ siehe Abschnitt 3 Tabelle 14

2 Material und Methoden

In diesem Abschnitt werden die in dieser Arbeit verwendeten Materialien aufgelistet und anschließend die angewandten Methoden der verschiedenen Analysen erläutert. Die Fragestellungen dieser Analysen wurden am Institut für Humangenetik der Christian-Albrechts-Universität zu Kiel bearbeitet, wenn nicht anders angegeben.

2.1 Material

Dieser Abschnitt der Arbeit beschreibt die Materialien, die für die verschiedenen Analysen verwendet wurden. Standardmäßig im Labor eingesetzte Geräte und Reagenzien werden im Folgenden nicht explizit erwähnt.

2.1.1 Probenmaterial, Kontrollen und Zelllinien

Im Folgenden werden das Probenmaterial sowie die verwendeten Kontrollen und Zelllinien der verschiedenen Analysen aufgeführt. Das Akquirieren des Probenmaterials von Patienten mit kongenitalen Herzfehlbildungen erfolgte an der Klinik für angeborene Herzfehler und Kinderkardiologie (UKSH, Campus Kiel), sowie über die Biomaterialbank des Kompetenznetzes Angeborene Herzfehler e.V..

2.1.1.1 Probenmaterial

Patienten¹² mit kongenitalen Herzfehlbildungen (*congenital heart defect*, CHD) bzw. deren Eltern wurden im Rahmen des stationären Aufenthalts bzw. ambulanter Betreuung in der Klinik für angeborene Herzfehler und Kinderkardiologie (UKSH, Campus Kiel) über die Zielsetzung der genetischen Studien, welche in Zusammenarbeit mit dem Kompetenznetz Angeborene Herzfehler e.V. durchgeführt wurden, aufgeklärt. Nach erfolgter Aufklärung der Patienten bzw. deren Eltern über die anonyme Verwendung der Proben für wissenschaftliche Analysen erfolgte die Probenentnahme inklusive der Erhebung relevanter klinischer Daten. Die im Rahmen der routinemäßig anstehenden Herzoperationen anfallenden **Herzgewebeproben**, welche in der Routine verworfen werden, konnten für die verschiedenen Analysen verwendet werden. Zusätzlich wurden verschiedene Herzgewebeproben *post mortem* im Rahmen der pathologischen Untersuchung entnommen. Insgesamt konnten 62 Herzgewebeproben von 47 Patienten mit verschiedenen CHDs, insbesondere mit der Diagnose des hypoplastischen Linksherzsyndroms (HLHS), akquiriert werden.

Es standen folgende klinische Angaben zu den Herzgewebeproben zur Verfügung: kardiale Diagnose, Gewebeentnahme-Art (Operationstechniken), Geschlecht, bekannte chromosomale Anomalien, Herzgewebetyp, Gewebeisolation und Alter bei Gewebeentnahme. In Tabelle 1 sind die klinischen Angaben sowie eine Auflistung der verschiedenen an den Proben durchgeführten Analysen aufgeführt. Bei allen Proben handelte es sich um frisch isoliertes Gewebe (fr), mit Ausnahme der *post mortem* (pm) entnommenen Gewebeproben. Die Probenentnahme und genaue topographische Zuordnung erfolgte durch die Kinderherzchirurgen der Herz- und Gefäßchirurgie des UKSH (Campus Kiel). Bei den durch das Kompetenznetz Angeborene Herzfehler e.V. zur Verfügung gestellten Proben (siehe Index 1 der Probenbezeichnung in Tabelle 1) erfolgte die Probenentnahme und topographische Zuordnung durch Herzchirurgen teilnehmender Kliniken.

¹² Im Rahmen dieses Textes schließt die männliche Bezeichnung stets die weibliche Bezeichnung mit ein.

Tabelle 1 Probenbezeichnungen und klinische Angaben zu den Herzgewebeproben bzw. Herzgewebe-DNA

Proben- Bezeichnung	Diagnose (HLHS-Subtyp)	Herz-OP/ Gewebe-entnahme	Geschlecht	CHR-Anomalie	Gewebe-tyt	Gewebe-Isolation	Alter [Tage]	Alter-Kat.	Analyse
HLHS_4	HLHS (Typ: MS/AS)	N	m		IAS	fr	4	J	M (P: 1, 2, 3, 5), E, R1
HLHS_8	HLHS (Typ: MS/AA)	N	w		IAS	fr	21	J	M (P: 1, 2, 3, 5), E, R1
HLHS_10	HLHS (Typ: MS/AS)	N	m		IAS	fr	3	J	M (P: 1, 2, 3, 5), R1
HLHS_11	HLHS (Typ: MA/AA)	H	m		IAS	fr	131	M	M (P: 1, 2, 3), R1
HLHS_12	HLHS (Typ: MA/AA)	N	m		IAS	fr	4	J	M (P: 1, 2, 3, 5), E, R1
HLHS_13	HLHS (Typ: MA/AA)	N	w		IAS	fr	4	J	M (P: 1, 2, 3, 5), E, R1
HLHS_15	HLHS (Typ: MA/AA)	N	w		IAS	fr	4	J	M (P: 1, 2, 3, 5), E, R1
HLHS_17	HLHS (Typ: MS/AA)	N	w		IAS	fr	4	J	M (P: 1, 2, 3, 5), R1
HLHS_20	HLHS (Typ: MS/AS)	N	m		IAS	fr	5	J	M (P: 1, 2, 3, 5), R1
HLHS_21*	HLHS (Typ: MS/AA)	H	m	*1	IAS	fr	135	M	M (P: 1, 2, 3), E, R1
HLHS_22*	HLHS (Typ: MS/AA)	H	m	*1	IAS	fr	141	M	M (P: 1, 2, 3), E, R1
HLHS_24	HLHS (Typ: MA/AA)	N	m		IAS	fr	5	J	M (P: 1, 2, 3, 5), R1
HLHS_29	HLHS (Typ: MA/AA)	H	m		IAS	fr	106	M	M (P: 1, 2, 3), R1
HLHS_35	HLHS (Typ: MA/AA)	N	w	*2	IAS	fr	2	J	M (P: 1, 2, 3, 5), E
HLHS_39	HLHS (Typ: MS/AA)	N	m		IAS	fr	6	J	M (P: 1, 2, 3, 5), R2
HLHS_64	HLHS (Typ: MS/AS)	N	m		IAS	fr	2	J	M (P: 1, 2, 3, 5), R2
HLHS_72	HLHS (Typ: MS/AS)	N	w		IAS	fr	9	J	M (P: 1, 2, 3, 5), R2
HLHS_76	HLHS (Typ: MA/AA)	N	m		IAS	fr	5	J	M (P: 1, 2, 3, 5), R2
HLHS_83	HLHS (Typ: MS/AA)	H	m		IAS	fr	81	M	M (P: 1, 2, 3), R2
HLHS_90	HLHS (Typ: MS/AS)	TCPC	m		IAS	fr	1429	A	M (P: 1, 2, 3)
HLHS_116	HLHS (Typ: MS/AA)	N	m		IAS	fr	44	M	M (P: 1, 2, 3), R2
HLHS_128	HLHS (Typ: MS/AS)	N	m		IAS	fr	6	J	M (P: 1, 2, 3, 5), R2
HLHS_129	HLHS (Typ: MA/AA)	H	w		IAS	fr	706	A	M (P: 1, 2, 3)
HLHS_130	HLHS (Typ: MS/AS)	N	w		IAS	fr	4	J	M (P: 1, 2, 3, 5), R2
HLHS_132	HLHS (Typ: MS/AS)	N	m		IAS	fr	10	J	M (P: 1, 2, 3, 5), E, R2
HLHS_157	HLHS (Typ: MS/AS)	N	m		IAS	fr	11	J	M (P: 1), R2
HLHS_182	HLHS (Typ: MA/AA)	TCPC	m		IAS, RA	fr	1151	A	M (P: 2)
HLHS_183	HLHS (Typ: MS/AS)	N	w		IAS	fr	3	J	R2
HLHS_195	HLHS (Typ: MS/AS)	N	m		IAS	fr	4	J	R2

Proben- Bezeichnung	Diagnose (HLHS-Subtyp)	Herz-OP/ Gewebe-entnahme	Geschlecht	CHR- Anomalie	Gewebe- typ	Gewebe- Isolation	Alter [Tage]	Alter- Kat.	Analyse
HLHS-TGA_117_LV	HLHS/TGA (Typ: MA/AA)	HTx	m		3xLV	fr	842	A	M (P: 2, 3)
HLHS-TGA_117_RV					3xRV				
HLHS-TGA_117_LA					3xLA				
HLHS-TGA_117_RA					3xRA				
Lvan_126_LV	LV-An	HTx	w		3xLV	fr	43	M	M (P: 2, 3)
Lvan_126_RV					2xRV				
Lvan_126_LA					3xLA				
Lvan_126_RA					3xRA				
DCM_175_LV	Fam. DCM	pm	m		LV	pm	4584	A	M (P: 2, 3, 4)
DCM_175_LA					LA				
DCM_175_RA					RA				
DCM_175_IAS					IAS				
DCM_175_Aorta					Aorta				M (P: 2, 4)
DCM_175_PA					PA				
DCM_175_AoV					AoV				
DCM_175_MV					MV				
DCM_175_TriV	TV								
AVkompl_141_RA	AVSD kompl.	AV	m	*3	RA	fr	148	M	M (P: 2, 3)
AVkompl_148_RA	AVSD kompl.	AV	m	*3	RA	fr	82	M	M (P: 2, 3)
AVkompl_96_RA	AVSD kompl.	AV	m	*3	RA	fr	117	M	M (P: 2, 3)
AVkompl_210_RA	AVSD kompl.	AV	w	*3	RA	fr	137	M	M (P: 2, 3)

Proben- Bezeichnung	Diagnose (HLHS-Subtyp)	Herz-OP/ Gewebe-entnahme	Geschlecht	CHR-Anomalie	Gewebe- typ	Gewebe- Isolation	Alter [Tage]	Alter- Kat.	Analyse
TA_55	TA, Dex.	H	m		IAS	fr	138	M	R2
TA_78	TA	N	m		IAS	fr	4	J	M (P: 1, 2, 3), R2
TA_85	TA	H	m		IAS	fr	344	A	M (P: 1, 2, 3), R2
TA_629 ¹	TA	TCPC	w		IAS	fr	5916	A	M (P: 1, 2, 3), R1
TA_187 ¹	TA	TCPC	m		IAS	fr	7180	A	M (P: 1, 2, 3), R1
TA_418 ¹	TA	GL	m		IAS	fr	89	M	M (P: 1, 2, 3), R1
TA_420 ¹	TA	TCPC	w		IAS	fr	1368	A	M (P: 1, 2, 3), R1
PA-VSD_330 ¹	PA, VSD	AP	w		IAS	fr	7	J	M (P: 1, 2, 3), R1
TA_631 ¹	TA	TCPC	m		IAS	fr	3091	A	M (P: 1, 2, 3)
TA_809 ¹	TA	AP	m		IAS	fr	10	J	M (P: 1, 2, 3), R2
TA_620 ¹	TA	AP	m		IAS	fr	4	J	M (P: 1, 2, 3), R2
F_FetalHerz	-	NA	w		GH	fr	40 SSW	fetal	M (P: 2, 5)
M_FetalHerz	-	NA	m		GH	fr	25 SSW	fetal	M (P: 2, 5)

Probenbezeichnung: jede Probennummer entspricht einem Patienten; HLHS_21/22*: monozygote Zwillinge; Index 1: vom Kompetenznetz Angeborene Herzfehler e.V. zur Verfügung gestellte Proben; grau schattiert: von diesem Gewebe stand lediglich die DNA-Probe zur Verfügung; **Diagnose:** HLHS: Hypoplastisches Linksherzsyndrom; TGA: Transposition der großen Arterien; LV-An: linksventrikuläres Aneurysma; fam. DCM: familiäre dilatative Kardiomyopathie; AVSD: atrioventrikulärer Septumdefekt; TA: Trikuspidalatresie; PA: Pulmonalatresie; Dex.: Dextrokardie; VSD: Ventrikelseptumdefekt. **HLHS Typ:** MA: Mitralatresie; MS: Mitralstenose; AA: Aortenatresie; AS: Aortenstenose. **Herz-OP (Operation)/ Gewebeentnahme** (Details siehe Abschnitt 8 Tabelle 30); N: Norwood; H: Hemifontan; TCPC: totale kavopulmonale Anastomose; GL: Glenn-Anastomose; AP: aortopulmonale Shunt-Anlage; HTx: Herztransplantation (Herzexplantat); AV: AV-Kanal-Korrektur; pm: *post mortem*; NA: keine Angabe; Geschlecht: m: männlich; w: weiblich. **CHR-Anomalien (chromosomale Anomalien):** *1: 46,XY,del(12)(p12.3),dup(X)(q26.2), *2: 46,XX,del(X)(p.11.4), *3: 47,XY,+21. **Gewebe**typ: IAS: interatriales Septum; LA: linkes Atrium; RA: rechtes Atrium; LV: linker Ventrikel; RV: rechter Ventrikel; MV: Mitralklappe; TriV/TV: Trikuspidalklappe; Aorta *ascendens* (Aorta); PA: Pulmonalarterie; AoV: Aortenklappe; GH: Gesamtherz (fetales Herzgewebe (Biocat (Heidelberg, Deutschland))); 2x/3x: Anzahl biologischer Replikate. **Gewebeisolation:** fr: frisch; pm: *post mortem*. **Alter: Alter bei Gewebeentnahme:** fetal: fetales Gewebe; SSW: Schwangerschaftswoche. **Alter-Kat.: Alterskategorie:** entsprechend Abschnitt 3.2.2.1 J = jung (≤ 3 bzw. 21 d); M = mittel (> 3 W bis ≤ 30 W bzw. 22 – 210 d); A = alt (> 30 W bzw. > 211 d). **Analyse:** M: DNA-Methylierungsanalyse; E: Gesamt-Exom-Sequenzierung; R1/R2: Gesamt-Transkriptom-Sequenzierung Analyse 1 bzw. 2.

Bis auf die im Zuge einer Herztransplantation (HTx) aus dem Herzexplantat entnommenen Gewebeproben der Patienten HLHS_TGA und LV-An, sowie die *post mortem* (pm) entnommenen Herzgewebeproben des DCM-Patienten, wurden alle weiteren Herzgewebeproben im Rahmen von Herzoperationen (Details siehe Abschnitt 8 Tabelle 30) entnommen. Die Herzgewebeprobenentnahme der hier aufgeführten Herzgewebeproben erfolgte im Mittel in einem Alter von 3 Jahren (Median = 85 Tage). Eine Gesamtübersicht über Patientenkollektiv und Herzgewebeproben mit Angaben zu dem Herzgewebetyp und dem mittleren Alter bei Gewebeenentnahme pro kardialer Diagnose ist in Abbildung 5 schematisch dargestellt.

CHD-Diagnosen und Gewebetyp

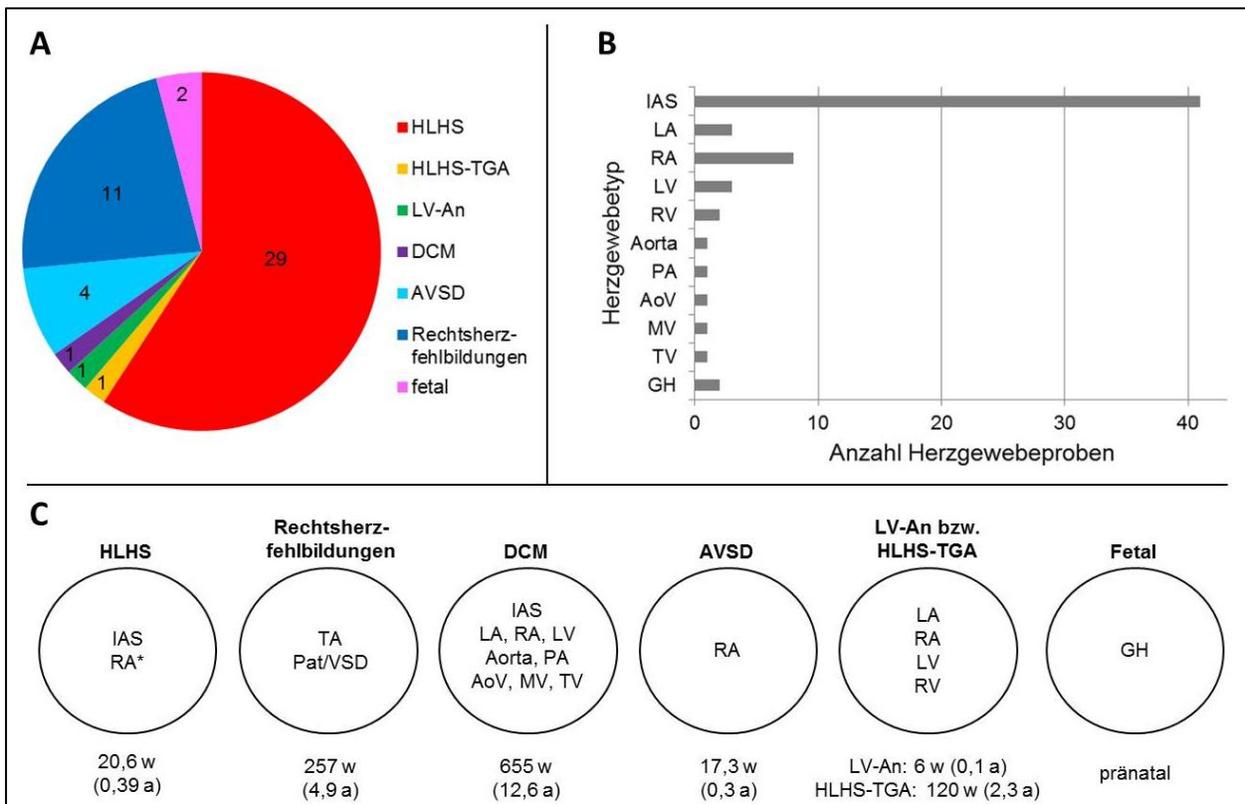


Abbildung 5 Gesamtübersicht über Herzgewebeproben und Patientenkollektiv. A: Übersicht der Anzahl von Patienten pro kardialer Diagnose der insgesamt 47 Patienten mit CHDs und der zwei fetalen Herzgewebe-Proben (Biocat (Heidelberg, Deutschland)). HLHS: Hypoplastisches Linksherzsyndrom; TGA: Transposition der großen Arterien; LV-An: linksventrikuläres Aneurysma; fam. DCM: familiäre dilatative Kardiomyopathie; AVSD: atrioventrikulärer Septumdefekt; Rechtsherzfehlbildungen: umfasst 10 Patienten mit Trikuspidalatresie (TA) und einen Patienten mit Pulmonalatresie und Ventrikelseptumdefekt (Pat/VSD)(Proben siehe Tabelle 1); Fetal: fetales Herzgewebe gesunder Individuen (Biocat). B: Übersicht über die Aufteilung nach Herzgewebetyp der insgesamt 64 analysierten Herzgewebeproben. IAS: interatriales Septum; LA: linkes Atrium; RA: rechtes Atrium; LV: linker Ventrikel; RV: rechter Ventrikel; PA: Pulmonalarterie; AoV: Aortenklappe; MV: Mitralklappe; TV: Trikuspidalklappe; GH: Gesamtherz (fetales Herzgewebe (Biocat)). C: Herzgewebetypen und mittleres Alter bei Herzgewebeprobenentnahme (w: Wochen, a: Jahre) pro kardialer Diagnose-Gruppe. RA*: von Patient HLHS_182 konnte sowohl IAS als auch RA gewonnen werden (siehe Tabelle 1).

Von insgesamt 29 **HLHS-Patienten** (Abbildung 5 A) konnte interatriales Septumgewebe (IAS) gewonnen werden, wobei zusätzlich von einem HLHS-Patienten, HLHS_182, auch rechtes Atriumgewebe (RA) gewonnen werden konnte (siehe Tabelle 1). Das mittlere Alter bei Herzgewebeentnahme ((m)AG) betrug 20,6 Wochen (Abbildung 5 C). Zwei der HLHS-Patienten im Probenkollektiv sind monozygote Zwillinge (siehe Tabelle 1 mit einem Sternchen gekennzeichnet). Bei letztgenannten Patienten konnten im Vorfeld der Arbeit chromosomale Anomalien festgestellt werden: eine Deletion in 12p12.3 (ca. 0,2 Mb) (gleicher Befund bei Vater) und eine Duplikation in Xq26.2 (ca. 0,3 Mb) (gleicher Befund bei Mutter) [247]. Ein weiterer HLHS-Patient hatte den genetischen Befund einer Deletion Xp.11.4 (siehe Tabelle 1). Bei allen HLHS-Patienten handelte es sich um isolierte nicht-syndromale (ohne Assoziation zu genetischen Syndromen bzw. extrakardialen Anomalien) Fälle. Nach Klassifikation der anatomischen Varianten des HLHS lagen innerhalb der 29 Patienten mit HLHS folgende Subtypen vor (siehe Tabelle 1, „Typ“):

- (1) Mitralstenose/Aortenstenose (MS/AS) mit 12/29 Patienten
- (2) Mitralstenose/Aortenatresie (MS/AA) mit 7/29 Patienten
- (3) Mitralatresie/Aortenatresie (MA/AA) mit 10/29 Patienten

Im Zuge von Herztransplantationen konnte aus dem Herzexplantat von einem Patienten mit **HLHS und Transposition der großen Arterien** (HLHS-TGA, AG: 2,3 Jahre), sowie von einem Patienten mit kongenitalem **linksventrikulärem Aneurysma** (LV-An, AG: 6 Wochen) je vier verschiedene Herzgewebetypen, d.h. linker/rechter Ventrikel (LV, RV) und linkes/rechtes Atrium (LA, RA), gewonnen werden (siehe Tabelle 1 und Abbildung 5 B, C). Von diesen vier Herzregionen konnten biologische Replikat exzidiert werden. Von einem Patienten mit **dilatativer Kardiomyopathie** (DCM, AG: 12,6 Jahre) stammten *post mortem* neun verschiedene Herzgewebeproben, d.h. IAS, LA, RA, LV, Aorta *ascendens* (Aorta), Pulmonalarterie (PA), Aortenklappe (AoV), Mitralklappe (MV) und Trikuspidalklappe (TV) (Abbildung 5 A, B, C). RA-Gewebeproben konnten von vier Patienten mit **atrioventrikulärem Septumdefekt** (AVSD, mAG: 17,3 Wochen) gewonnen werden. Diese Patienten wiesen als genetischen Befund eine Trisomie 21 auf. Zum Kollektiv der **Rechtsherzfehlbildungen** mit insgesamt 11 IAS-Gewebe-Proben zählten Patienten mit Trikuspidalatresie (TA) und ein Patient mit Pulmonalatresie (PA_t) und Ventrikelseptumdefekt (VSD). Das mAG der Patienten mit Rechtsherzfehlbildungen betrug 4,9 Jahre (Abbildung 5 C). Die Proben TA_629, TA_187, TA_418, TA_420, PA-VSD_330, TA_631, TA_809 und TA_620 aus diesem Kollektiv (siehe Index 1 der Probenbezeichnung in Tabelle 1) wurden vom Kompetenznetz Angeborene Herzfehler e.V. zur Verfügung gestellt.

Um zusätzlich Herzgewebematerial aus kardial gesundem Gewebe analysieren zu können, wurde DNA aus **fetalem Herzgewebe** (BioCat (Heidelberg, Deutschland)) verwendet („Fetal“, Abbildung 5). Hierbei handelt es sich um kommerziell erhältliche DNA aus „Gesamtherz“ („GH“, Angabe BioCat), jeweils von einem gesunden weiblichen (40. Woche) und männlichen Donor (25. Woche).

Insgesamt standen somit 62 Herzgewebeproben von 47 Patienten mit verschiedenen CHDs, sowie fetale „GH“-Proben zweier gesunder Individuen zur Verfügung. Dies teilte sich auf in 41 IAS-, drei LA-, acht RA-, drei LV-, zwei RV-, und je eine Aorta-, PA-, AoV-, MV-, TV-Gewebeprobe, sowie die beiden fetalen „GH“-Proben (Abbildung 5 B).

Histologie Herzgewebeproben

Mit Ausnahme der Herzgefäß- (Aorta, PA) und Herzklappenstrukturen (AoV, MV, TV) stammen alle weiteren Herzgewebeproben aus dem Myokard (Herzmuskelgewebe) [248]. Die Proben aus dem IAS und die Atriumproben besitzen zusätzlich einen Endokardanteil [248]. Die Herzklappen bestehen aus Einfaltungen des Endokards [248]. Bei der Gewebeexzision, die *post mortem* am Herzen des DCM-Patienten durchgeführt wurde, erfolgte im Falle der Herzklappenstrukturen eine Gewebeentnahme im Bereich des Aortenklappenrings (AoV-Probe) sowie im Bereich der Klappenaufhängung bei der MV und der TV. Der Aortenklappenring enthält kollagenes Bindegewebe [248]. Da Papillarmuskeln ein Bestandteil der Klappenaufhängung darstellen [249], enthalten die Proben der MV und TV einen Muskelgewebsanteil. Die Gefäßstrukturen Aorta *ascendens* und Pulmonalarterie bestehen hauptsächlich aus Endothel und Bindegewebe, sowie aus einer Schicht glatter Muskulatur [248].

Des Weiteren wurde **peripheres Blut** (pB) von Patienten mit HLHS und deren Eltern sowie von Familien mit Catel-Manzke-Syndrom entnommen. Die Eltern der analysierten HLHS-Patienten waren jeweils phänotypisch gesund und wiesen keine kardiale Diagnose auf. Bei minderjährigen Patienten wurden Blutentnahmen ausschließlich im Rahmen medizinisch-indizierter Routineuntersuchungen, Herzkatheteruntersuchungen oder Operationen durchgeführt. Einige DNA¹³-Proben, extrahiert aus pB, wurden vom Kompetenznetz Angeborene Herzfehler e.V. zur Verfügung gestellt. Die jeweiligen relevanten klinischen Angaben zu den Blutproben sind in Tabelle 2 zusammengefasst.

Tabelle 2 Probenbezeichnungen und klinische Angaben der pB-Proben

Probenbezeichnung	Familien-Zuordnung	Diagnose	Geschlecht
D12/579	Familie_1_Index	CMS	m
D12/581	Familie_1_Sohn	normal	m
D12/775	Familie_2_Index	CMS	w
D12/776	Familie_2_Mutter	normal	w
D11/527	Familie_3_Index	CMS (mit VSD)	w
D11/528	Familie_3_Mutter	normal	w
D11/1087	Familie_3_Vater	normal	m
HLHS_4_Index (pB-Probe zu HLHS_4)	Familie_4_Index	HLHS	m
HLHS_4_Mutter	Familie_4_Mutter	normal	w
HLHS_4_Vater	Familie_4_Vater	normal	m
HLHS_130_Index-1 (pB-Probe zu HLHS_130)	Familie_5_Index-1	HLHS	w
HLHS_130_Index-2	Familie_5_Index-2	HLHS	m
HLHS_130_Bruder	Familie_5_Bruder	normal	m
HLHS_130_Mutter	Familie_5_Mutter	normal	w
HLHS_130_Vater	Familie_5_Vater	normal	m
HLHS_132_Index (pB-Probe zu HLHS_132)	Familie_6_Index	HLHS	m
HLHS_132_Mutter	Familie_6_Mutter	normal	w
HLHS_132_Vater	Familie_6_Vater	normal	m
Index_Fam7	Familie_7_Index	HLHS	w
Mutter_Fam7	Familie_7_Mutter	normal	w
Vater_Fam7	Familie_7_Vater	normal	m
HLHS_21_pB	Familie_8_Index-1	HLHS	m
HLHS_22_pB	Familie_8_Index-2	HLHS	m
Mutter_Fam8	Familie_8_Mutter	normal	w
Vater_Fam8	Familie_8_Vater	normal	m

CMS: Catel-Manzke-Syndrom; HLHS: Hypoplastisches Linksherzsyndrom; VSD: Ventrikelseptumdefekt; m: männlich; w: weiblich.

¹³ Die Bezeichnung „DNA“ umfasst in dieser Arbeit grundsätzlich die genomische DNA (gDNA), wenn nicht anders angegeben.

2.1.1.1.1 Probenentnahme und Lagerung

Das Akquirieren der Herzgewebeproben von Patienten mit CHDs erfolgte größtenteils an der Klinik für angeborene Herzfehler und Kinderkardiologie (UKSH, Campus Kiel), wenn nicht anders angegeben. Für das System der Probenentnahme und Lagerung wurden vom Kompetenznetz Angeborene Herzfehler e.V. für diese Zwecke zur Verfügung gestellte Kryo-Reagenzgefäße, EDTA-Blut-Monovetten und Probenbegleitscheine verwendet. In Abbildung 6 ist schematisch die Probenentnahme, Lagerung und anschließende Extraktion von genomischer DNA (gDNA) bzw. RNA dargestellt.

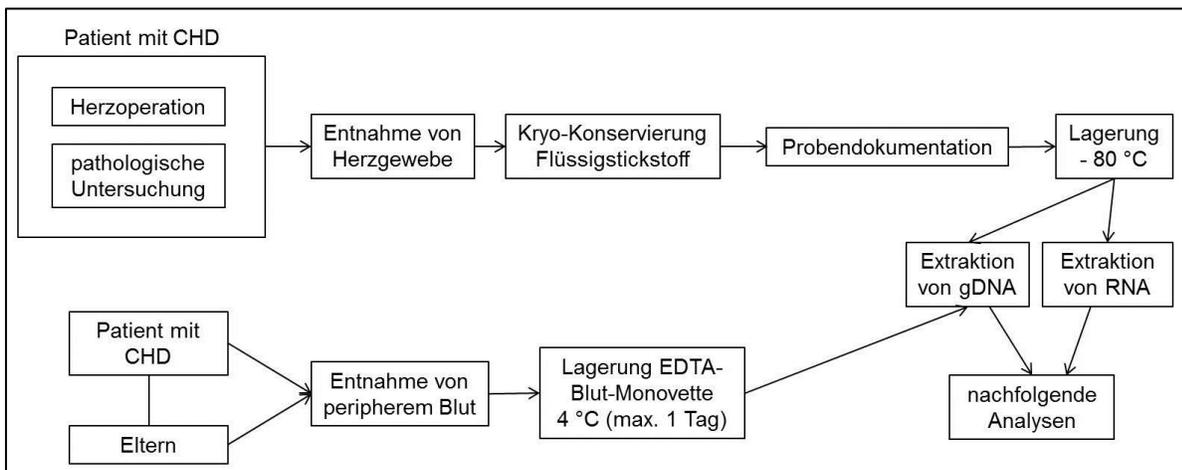


Abbildung 6 Schematische Darstellung der Probenentnahme, Lagerung und anschließenden Extraktion von gDNA bzw. RNA. CHD: kongenitale Herzfehlbildung

Sobald eine **Herzgewebeprobe** im Rahmen einer Herzoperation bzw. einer pathologischen Untersuchung (hier 12 h nach Todeseintritt) entnommen wurde, konnte diese in Farb- und Barcode-markierte Kryo-Reagenzgefäße überführt und in flüssigem Stickstoff schockgefroren werden. Um eine kurze Zeitspanne zwischen Probenentnahme und Einfrieren zu gewährleisten, wurde ein tragbarer Stickstoffbehälter in einem Vorraum des Operationssaals genutzt. Die Proben wurden hierbei unter sterilen Bedingungen innerhalb von einer Minute eingefroren, um eine Degradation des Probenmaterials auf ein Minimum zu reduzieren. Nach erfolgter Probendokumentation mittels Probenbegleitschein wurden diese bei – 80 °C gelagert.

Zur Entnahme der **pB-Proben** bei Patienten mit CHDs und deren Eltern wurden EDTA-Blut-Monovetten (Sarstedt S-Monovette® 4,9 ml/9 ml) verwendet und diese anschließend bis zur Extraktion von gDNA max. einen Tag bei 4° C gelagert.

2.1.1.2 Kontrollen und Zelllinien

Für die verschiedenen Analysen wurden stets Kontrollen eingesetzt, um technisch valide Durchführungen der verschiedenen Labormethoden zu sichern und eine kontrollierte Auswertung der Ergebnisse zu gewährleisten.

Für **PCR-Reaktionen** (*polymerase chain reaction* (Polymerase-Kettenreaktion)) wurden als Positivkontrollen sogenannte DNA-Normalkontrollen, d.h. unauffälliges DNA-Probenmaterial aus der Routinediagnostik des Instituts für Humangenetik in Kiel, verwendet (Tabelle 3). Diese DNA-Normalkontrollen dienten auch der Austestung und Optimierung von PCR-Bedingungen. Als Kontaminationskontrolle diente bidestilliertes Wasser (Aqua bidest (ddH₂O)). Die Methode der **Bisulfit-Pyrosequenzierung** (BSPS) erforderte fünf verschiedene Kontrollen (Tabelle 3), die in Abschnitt 2.2.2.3 im Detail aufgeführt werden. Die verwendeten Kontrollen für die PCR-basierte DNA-Methylierungsanalyse mittels **MassARRAY-EpiTYPER-Technik** (MassARRAY® EpiTYPER®, Sequenom Inc.; im Folgenden EpiTYPER-Technik abgekürzt) werden in Abschnitt 2.2.2.4 genauer beschrieben. Für die **quantitative Real-Time PCR** wurde als Positivkontrolle die kommerziell erhältliche Herz-cDNA (*complementary DNA* (komplementäre DNA)) von Clontech Laboratories Inc. (Mountain View, USA), als Negativkontrolle eine DNA-Normalkontrolle und ddH₂O als Kontaminationskontrolle verwendet (Tabelle 3).

Zelllinien-DNA, die aus Burkitt-Lymphom-Zelllinien (Leibniz Institut Deutsche Sammlung von Mikroorganismen und Zellkulturen (DSMZ; Braunschweig, Deutschland)) extrahiert vorlag, wurde von der Kollegin Frau Dr. rer. nat. Rabea Wagener zur Verfügung gestellt. Die Zelllinien-DNA wurde zur Austestung und Optimierung der PCR-Reaktionen für die anschließende Bisulfit-Pyrosequenzierung (siehe 2.2.2.4) verwendet. In Tabelle 3 sind die verwendeten DNA-Kontrollen und Zelllinien-DNA sowie die dazugehörigen Labortechniken zusammengefasst.

Tabelle 3 Kontrollen und Zelllinien-DNA, Anwendung und Herkunft

DNA-Material/Kontrollen	Herkunft DNA-Material bzw. Hersteller	Anwendung/Labortechnik
DNA-Normalkontrolle	unauffälliges DNA-Material aus der Routinediagnostik des Instituts für Humangenetik, Kiel	BSPS PCR für Sequenzierungen quantitative Real-Time PCR
Zelllinien-DNA Raji	Burkitt-Lymphom-Zelllinie Raji [250], Zellkultur des Instituts für Humangenetik, Kiel	BSPS
Zelllinien DNA Ramos	Burkitt-Lymphom-Zelllinie Ramos (n.p.), Zellkultur des Instituts für Humangenetik, Kiel	BSPS
hoch-methylierte Kontroll-DNA (> 85%)	EpigenDx, Hopkinton, USA	BSPS, Epityper-Technik
unmethylierte Kontroll-DNA (< 5 %)	EpigenDx, Hopkinton, USA	Epityper-Technik
Gesamt-Genom-amplifizierte DNA (WGA, <i>Whole Genome Amplification</i>)	Gesamt-Genom-Amplifizierte DNA aus Pool DNA	BSPS
Pool-DNA	DNA-Pool von zehn männlichen und zehn weiblichen gesunden, unauffälligen Individuen aus der Routinediagnostik des Instituts für Humangenetik, Kiel	BSPS
Herz-cDNA	Clontech Laboratories Inc., Mountain View, USA	quantitative Real-Time PCR
bidestilliertes Wasser (Aqua bidest (ddH ₂ O))	Fresenius Kabi AG, Bad Homburg, Deutschland	PCR für Sequenzierungen quantitative Real-Time PCR BSPS, Epityper-Technik

DNA-Material: Bezeichnung von DNA- und cDNA-Material; Herkunft DNA-Material: Angaben über Herkunft, Firmen (Firmenname, Ort, Staat); BSPS: Bisulfit-Pyrosequenzierung; n.p.: nicht publiziert; Anwendung/Labortechnik: Anwendungsbereich des DNA-Materials. In dieser Arbeit erfolgt lediglich bei der ersten Erwähnung der DNA-Materialien im Fließtext eine Angabe zur Herkunft bzw. Hersteller. Weitere Angaben können aus dieser Tabelle entnommen werden und werden daher im Fließtext nicht explizit erwähnt.

2.1.2 Material für Laborarbeit

In diesem Abschnitt wird das Material für die Laborarbeit beschrieben. Hierbei handelt es sich um Primer für die verschiedenen Labormethoden (siehe 2.1.2.1), verwendete Chemikalien und Reagenzien (siehe 2.1.2.2), Puffer und Lösungen (siehe 2.1.2.3), Kits (siehe 2.1.2.4), sowie verwendete Geräte und Zubehör (siehe 2.1.2.5). Die für die Labormethoden notwendigen Geräte standen im Institut für Humangenetik der Christian-Albrechts-Universität zu Kiel zur Verfügung, sofern nicht anders angegeben.

2.1.2.1 Primer

Alle Primer wurden grundsätzlich im Anschluss an das Primer-Design mittels UCSC BLAT-Analyse (siehe 2.1.3.2) gegen die Genom-Referenz GRCh37/hg19 hinsichtlich der Spezifität überprüft. Bezüglich SNPs (*single nucleotide polymorphism* (Einzel-Nukleotid-Polymorphismen)) in der Primer-Sequenz wurde mittels UCSC Genome Bioinformatics Datenbank [251] (siehe 2.1.3.2) analysiert und ggf. ein neues Primer-Design durchgeführt. Die

Oligonukleotide für die Sequenzierung von PCR-Produkten und für die BSPS-Methode wurden als Lyophilisate von Biomers (Ulm, Deutschland) erhalten und in ddH₂O gelöst (final: 200 pmol/μl). Die Oligonukleotide für die Epityper-Technik wurden in gelöster Form (100 pmol/μl) von Metabion (München, Deutschland) erhalten. In Abschnitt 8 sind die Primer für die Sequenzierung von PCR-Produkten (Tabelle 31), sowie die Primer für die Methode der BSPS (Tabelle 32) aufgelistet. Die Primer für die Epityper-Technik werden in Tabelle 4 aufgeführt.

Tabelle 4 Primer für die Epityper-Technik

Primer Name [TargetID_HGNC_fw/rv]	Primer Sequenz [5'Sequenz-Tag...3']	Position CHR (GRCh37/hg19) [CHR:Start-Ende]	Ampl.- Länge [bp]	Primer T _m [°C]
cg17710779_FGFR2_fw	aggaagagagTATTTGAAAAGTTTGTTTAGTTT GAATAGA	chr10:123356140- 123356438	299	60
cg17710779_FGFR2_rv	<u>cagtaatacgcactcactatagg</u> gagaaggctAAAAAATT CCTCAACAAATTTAAATCTC			
cg17774559_IRX4_fw	aggaagagagTTTTATGTTTATGATGTTTGGG GTT	chr5:1879556- 1880012	457	60
cg17774559_IRX4_rv	<u>cagtaatacgcactcactatagg</u> gagaaggctCAAAAAAA ACACTATACCTACTACTCCA			

Primer Name [TargetID_HGNC_fw/rv]: [TargetID]: Ziel-CpG, zu analysierendes CG-Dinukleotid mit Angabe der Illumina TargetID; HGNC: Genname nach HUGO Gene Nomenclature Committee (HGNC) [265] (siehe 2.1.3.2); fw: Vorwärts-Primer, rv: Rückwärts-Primer. Primer Sequenz [5'Sequenz-Tag...3']: Primersequenz in 5' nach 3'-Richtung mit 5'Sequenz-Tag (Tag in kleinen Buchstaben, T7-Promotor unterstrichen dargestellt). Position CHR (GRCh37/hg19) [CHR:Start-Ende]: chromosomale Position des Amplifikats nach GRCh37/hg19. Ampl.-Länge [bp]: Amplifikatlänge in Basenpaaren (bp). Primer T_m [°C]: Primer-Schmelztemperatur (T_m) in °C.

In Tabelle 5 werden die Primer für die quantitative Real-Time PCR aufgelistet. Die Oligonukleotide für die quantitative Real-Time PCR wurden als kommerzielle Primer-Assays von Qiagen (Hilden, Deutschland) erworben.

Tabelle 5 Primer für die quantitative Real-Time PCR

Gen (HGNC)	detektiertes Transkript	Primer Name [HGNC_Exone_fw/rv]	Primer Sequenz [5'...3']	Position CHR (GRCh37/hg19) [CHR:Start-Ende]	Ampl.- Länge [bp]
PDGFRA	NM_006206	PDGFRA_15-16_fw	CATGCTAGAAAGGAAAG AGGTTTCT	chr4:55146558- 55151577	132
		PDGFRA_15-16_rv	TCTGAGTTATCATCTGAA AGGAGGT		

Gen (HGNC): Genname nach HUGO Gene Nomenclature Committee (HGNC); detektiertes Transkript: NM-Nummer nach NCBI; Primer Name [HGNC_Exone_fw/rv]: Genname (HGNC), Exonnummern, die das Amplifikat überspannt, fw: Vorwärts-Primer, rv: Rückwärts-Primer; Primer-Sequenz [5'...3']: Primersequenz in 5' nach 3' Richtung; Position CHR (GRCh37/hg19) [CHR:Start-Ende]: chromosomale Position des Amplifikats nach GRCh37/hg19; Ampl.-Länge [bp]: Amplifikatlänge in Basenpaaren (bp).

2.1.2.2 Chemikalien und Reagenzien

In Tabelle 6 werden die verwendeten Chemikalien und Reagenzien aufgeführt.

Tabelle 6 Auflistung verwendeter Chemikalien und Reagenzien und deren Hersteller

Chemikalien und Reagenzien	Hersteller
10 % Tween-20	Thermo Scientific, Waltham, USA
10x Tris-Borat-ETDA Puffer (TBE Puffer)	Life Technologies, Carlsband, USA
5 M NaCl	Life Technologies, Carlsband, USA
Agarose	Biozym, Hessisch Oldendorf, Deutschland
Aqua bidest (ddH ₂ O)	Fresenius Kabi AG, Bad Homburg, Deutschland
Betain-Lösung 5 M	Sigma Aldrich, St. Louis, USA
Bromphenolblau	Serva, Heidelberg, Germany
Denhardt's Solution 50x Concentrate	Sigma Aldrich, St. Louis, USA
Dimethylsulfoxid (DMSO) ReagentPlus ^R ≥ 99,5 %	Sigma Aldrich, St. Louis, USA
dNTP (Desoxynukleosidtriphosphat) Set 4 x 25 µmol	Roche, Rotkreuz, Schweiz
EB-Buffer	Qiagen, Hilden, Deutschland
Ethanol absolut	Th. Geyer, Renningen, Deutschland
Ethidiumbromid	Sigma Aldrich, St. Louis, USA
Ethylenediaminetetraacetic acid disodium salt solution (EDTA)	Life Technologies, Carlsband, USA
Formamid ≥ 99,5%	Sigma Aldrich, St. Louis, USA
GeneChip® Fragmentation Reagent	Affymetrix, Santa Clara, USA
GeneChip® DNA Labeling Reagent	Affymetrix, Santa Clara, USA
Glyzerin	Sigma Aldrich, St. Louis, USA
Hering Sperma DNA	Promega, Madison, USA
Human Cot-1 DNA	Life Technologies, Carlsband, USA
HyperLadder TM 1kb	Bioline, Taunton, USA
Isopropanol	Sigma Aldrich, St. Louis, USA
MES Hydrat ≥ 99,5 %, pH-,5-40 SigmaUltra	Sigma Aldrich, St. Louis, USA
MES Natriumsalz ≥ 99,5 %,	Sigma Aldrich, St. Louis, USA
Molecular Biology Grade Water	Lonza, Köln, Deutschland
Natriumhydroxid	Merck Millipore, Darmstadt, Deutschland
POP-7 TM -Polymer (3130 Genetic Analyzer)	Life Technologies, Carlsband, USA
Puffer 10x mit EDTA	Life Technologies, Carlsband, USA
PyroMark Annealing Buffer	Qiagen, Hilden, Deutschland
PyroMark Binding Buffer	Qiagen, Hilden, Deutschland
Sequencing Separation Buffer	Beckmann Coulter, Krefeld, Deutschland
Streptavidin Sepharose [®] High Performance	GE Healthcare, München, Deutschland
Streptavidin, R-phycoerythrin Conjugate	Life Technologies, Carlsband, USA
Terminal Deoxynucleotidyl Transferase	Affymetrix, Santa Clara, USA
TriDye TM 100 bp DNA Ladder	New England Biolabs, Frankfurt Main, Deutschland
Tris(hydroxymethyl)-aminomethan (TRIS)	Merck Millipore, Darmstadt, Deutschland
Ultra Pure TM 20x SSPE Puffer	Life Technologies, Carlsband, USA

In dieser Arbeit erfolgt lediglich bei der ersten Erwähnung dieser Chemikalien bzw. Reagenzien im Fließtext eine Angabe zum Hersteller. Weitere Angaben zum Firmensitz der Hersteller können aus dieser Tabelle entnommen werden und werden daher im Fließtext nicht explizit erwähnt.

2.1.2.3 Puffer und Lösungen

In Tabelle 7 werden die in der Arbeit verwendeten Puffer und Lösungen, sowie deren chemische Bestandteile aufgelistet.

Tabelle 7 Puffer und Lösungen, sowie deren chemische Bestandteile

Puffer/Lösung (Anwendung)	Bestandteile
0,1 M Natriumhydroxidlösung (450K-Array)	1 M Stammlösung (2 g NaOH add ddH ₂ O auf 50 ml) 1:10 mit ddH ₂ O verdünnt
0,2 M Natriumhydroxid - Denaturierungslösung (BSPS)	2 M Stammlösung (80 g NaOH add ddH ₂ O auf 1 l) 1:10 mit ddH ₂ O verdünnt
0,5 % EDTA-Lösung (450K-Array)	7,3 g EDTA add ddH ₂ O auf 50 ml
70 % Ethanol	140 ml 99 % Ethanol add ddH ₂ O auf 200 ml
95 % Formamid/1 mM EDTA-Lösung (450K-Array)	38 ml 100 % Formamid 80 µl 0,5 M EDTA add ddH ₂ O auf 40 ml
Agarose-Gel-Ladepuffer (Agarose-Gelelektrophorese)	1,9 mM Bromphenolblau 3,1 M Glycerin add ddH ₂ O auf 100 ml Lagerung: 4°C
Anti-Streptavidin Antikörperlösung 0,5 mg/ml (SNP-Array)	0,5 mg Anti-Streptavidin-Antikörper add ddH ₂ O auf 1 ml
Färbe-Puffer (SNP-Array)	360 µl Ultra Pure™ 20x SSPE Puffer 3,96 µl Tween-20 (3 %) 24 µl 50 x Denhardt's Solution 800 µl Molecular Biology Grade Water
Holding-Puffer (SNP-Array)	50 µl 10 % Tween 20 4,15 ml MES Stammlösung 9,25 ml NaCl (5 M) 36,55 ml ddH ₂ O
N-Morpholinoethansulfonsäure (MES) Stammlösung (SNP-Array)	35,2 g MES Hydrat 96,65 g MES Natriumsalz add 500 ml mit Molecular Biology Grade Water pH-Einstellung: 5,6 – 7,5 Sterilfiltrierung mittels 0,2 µm Filter
TBE, 1x (Agarose-Gelelektrophorese)	100 ml 10x TBE Puffer add ddH ₂ O auf 1 l
Waschpuffer (BSPS)	6,5 g TRIS add ddH ₂ O auf 5 l
Waschpuffer A (SNP-Array)	1 ml 10 % Tween 20 300 ml Ultra Pure™ 20x SSPE Puffer add ddH ₂ O auf 1 l Sterilfiltrierung mittels 0,2 µm Filter
Waschpuffer B (SNP-Array)	1 ml 10 % Tween 20 30 ml Ultra Pure™ 20x SSPE Puffer Add ddH ₂ O auf 1 l Sterilfiltrierung mittels 0,2 µm Filter

add: addiere; BSPS: Bisulfit-Pyrosequenzierung; 450K-Array: Infinium HumanMethylation450 BeadChip Array (Illumina); SNP-Array: Genome-Wide Human SNP Array 6.0 (Affymetrix). Die aufgelisteten Bestandteile beziehen sich teilweise auf Angaben in Tabelle 6.

2.1.2.4 Kits

In Tabelle 8 werden die für die verschiedenen Labortechniken angewendeten Kits sowie deren Hersteller aufgelistet.

Tabelle 8 Kits und Hersteller

Kit Bezeichnung	Hersteller
AccuPrime™ Polymerase System	Life Technologies, Carlsband, USA
Agencourt® AMPure® XP Kit	Beckmann Coulter, Krefeld, Deutschland
AmpliTag Gold® Polymerase System	Life Technologies, Carlsband, USA
Big Dye Terminator® v1.1 Cycle Sequencing Kit	Applied Biosystems, Foster City, USA
EpiTYPER® T Complete Reagent Set	Sequenom Inc., San Diago, USA
EZ DNA Methylation™ Kit	Zymo Research, Freiburg, Deutschland
Genome Wide Human SNP Nsp/Sty Assay Kit 6.0	Affymetrix, Santa Clara, USA
Genome-Wide Human SNP 6.0 Array Kit	Affymetrix, Santa Clara, USA
Gentra® Puregene® Blood Kit	Qiagen, Hilden, Deutschland
Illustra GenomiPhi™ V2 DNA Amplification Kit	GE Healthcare, München, Deutschland
Infinium HumanMethylation450 BeadChip Kit	Illumina, San Diego, USA
Montage™ SEQ96 Sequencing Reaction Cleanup Kit	Merck Millipore, Darmstadt, Germany
PyroMark® Gold 96 Reagents Kit	Qiagen, Hilden, Deutschland
PyroMark® PCR Kit	Qiagen, Hilden, Deutschland
Qiagen MinElute® 96 UF PCR Purification Kit	Qiagen, Hilden, Germany
QuantiTect® Reverse Transcription Kit	Qiagen, Hilden, Germany
QuantiTect® SYBR®Green PCR Kit	Qiagen, Hilden, Germany
Qubit® dsDNA BR Assay Kit	Life Technologies, Carlsband, USA
Qubit® dsDNA HS Assay Kit	Life Technologies, Carlsband, USA
T4 DNA Ligase (400 U/µl)	New England Biolabs, Frankfurt Main, Deutschland
TITANUM® Taq Polymerase Kit	Clontech Laboratories Inc., Mountain View, USA

In dieser Arbeit erfolgt bei Erwähnung dieser Kits im Fließtext eine Angabe zum Hersteller. Weitere Angaben zum Firmensitz der Hersteller können aus dieser Tabelle entnommen werden und werden daher im Fließtext nicht explizit erwähnt.

2.1.2.5 Geräte und Zubehör

In Tabelle 9 werden die für die verschiedenen Labormethoden verwendeten Geräte und das jeweilige Zubehör aufgeführt.

Tabelle 9 Geräte und Zubehör

Gerät/Zubehör (Anwendung)	Typbezeichnung	Hersteller
0,2 µm Filter	Filter Steritop-GP	Millipore, Darmstadt, Deutschland
4-Kapillarsequenzierer (PCR-basierte DNA-Sequenzierung)	ABI Prism®3130 Genetic Analyzer	Life Technologies, Carlsband, USA
Agarose-Gel-Dokumentationsanlage	BioDoc Analyzer	Biometra, Göttingen, Deutschland
Agarose-Gel-Elektrophoresekammersystem	peqLab	peqLab Biotechnologie, Erlangen, Deutschland
Aufreinigungsstation (WGA-Herstellung)	Wizard® DNA Clean-Up System	Promega, Madison, USA
Exsikator (450K-Array)	Bel-art	Illumina, San Diego, USA
Filtersonden für Aufreinigungsstation (BSPS)	PyroMark Vakuum Prep Filter Probe	Qiagen, Hilden, Deutschland
Fluorometer	Qubit 2.0 Fluorometer	Life Technologies, Carlsband, USA
Heizplatte mit Aufsatz für 96-Well Platte (BSPS)	Hotplate	VWR International GmbH, Darmstadt, Deutschland

Gerät/Zubehör (Anwendung)	Typbezeichnung	Hersteller
Hybridisierungsöfen (450K-Array)	Hybrid Oven	Illumina, San Diego, USA
Hybridisierungsöfen (SNP-Array)	GeneChip Hybridisierungsöfen 640	Affymetrix, Santa Clara, USA
Inkubator (450K-Array)	Rack-Chamber, 48Pos, Te-Flow, Tecan	Illumina, San Diego, USA
Kartusche für Nucleotide, Substrat- und Enzym-Mix (BSPS)	PSQ 96 Reagent Cartridge	Qiagen, Hilden, Deutschland
Massenspektrometer für Epityper-Technik	MassARRAY™ Analyzer Compact	Sequenom Inc., San Diego, USA
Mikrozentrifuge	Model SD	Carl ROTH, Karlsruhe, Deutschland
Nanodispenser zur Beladung von SpectroCHIP® Arrays	MassARRAY™ Nanodispenser RS1000	Sequenom Inc, San Diego, USA
PCR-Aufreinigungsstation (BSPS)	PyroMark® Q96 Vacuum Workstation	Qiagen, Hilden, Deutschland
Platteninkubator (450K-Array)	Hybex® Microsample Incubator	SciGene, Sunnyvale, USA
Plattenschüttler (450K-Array)	High-Speed Microplate Shaker	Illumina, San Diego, USA
Plattenschüttler (BSPS/ PCR-basierte DNA-Sequenzierung)	IKA® MS3 basic	IKA, Wilmington, USA
Pyrosequenziergerät	PyroMark Q96 ID	Qiagen, Hilden, Deutschland
Real-Time PCR Instrument	LightCycler® 480 II	Roche, Rotkreuz, Switzerland
Scanner (450K-Array)	iScan	Illumina, San Diego, USA
Scanner (SNP-Array)	GeneChip® Scanner 3000	Affymetrix, Santa Clara, USA
Stromversorgung der Agarose-Gelkammern	El.pho.-Powerpack P24	Life Technologies, Carlsbad, USA
Stromversorgung der Agarose-Gelkammern	Electrophoresis Power Supply ST606	Life Technologies, Carlsbad, USA
Thermoblock	UNO-Thermoblock	Biometra, Göttingen, Deutschland
Thermocycler	T3000 Thermocycler	Biometra, Göttingen, Deutschland
Thermocycler	DNA Engine-Thermocycler	Bio-Rad, München, Deutschland
Thermocycler für Epityper-Technik	Applied Biosystems® GeneAmp® PCR System 9700	Life Technologies, Carlsbad, USA
Thermomixer	Thermomixer Compact	Eppendorf, Hamburg, Deutschland
Tischzentrifuge	Mikro200	Hettich, Tuttlingen, Deutschland
Vakuumpumpe (450K-Array)	Model 2534C-02	Welch, Niles, USA
Vakuumpumpe (BSPS)	MultiScreen HTS Vakuum Manifold	Millipore, Darmstadt, Deutschland
Vakuumpumpe (Herstellung von WGA-Kontroll-DNA)	Labport	KNF, Trenton, USA
Vakuumpumpe (PCR-basierte DNA-Sequenzierung)	Millivac™ Vacuum Pump	Millipore, Darmstadt, Deutschland
Verschweißgerät für 96-Well Platten	ALPS™ 25	Life Technologies, Carlsbad, USA
Vortexgerät	REAX 2000	Heidolph, Schwabach, Deutschland
Vortexgerät	Vortex Genie 2	Carl ROTH, Karlsruhe, Deutschland

Gerät/Zubehör (Anwendung)	Typbezeichnung	Hersteller
Wasch- und Färbe-Station (SNP-Array)	GeneChip® Fluidics Station 450	Affymetrix, Santa Clara, USA
Zentrifuge	Megafuge 1.OR	Heraeus, Hanau, Deutschland
Zentrifuge	Biofuge primoR	Heraeus, Hanau, Deutschland
Zentrifuge	Multifuge X1R	Heraeus, Hanau, Deutschland
Zentrifuge (Platte, 450K-Array)	Megafuge 40R (Rotor: 75003607)	Heraeus Kulz GmbH, Dormagen, Deutschland
Zentrifuge für 1,5 ml Reaktionsgefäß	Mikro 200 (Rotor: 1195-A)	Hettich, Tuttlingen, Germany
Zentrifuge für 1,5 ml Reaktionsgefäß (DNA-Extraktion)	Biofuge primo R (Rotor: 7593)	Heraeus, Hanau, Germany

450K-Array: Infinium HumanMethylation450 BeadChip Array; BSPS: Bisulfit-Pyrosequenzierung; SNP-Array: Genome-Wide Human SNP Array 6.0. In dieser Arbeit erfolgt lediglich bei der ersten Erwähnung dieser Geräte im Fließtext eine Angabe zur Typbezeichnung und Hersteller. Weitere Angaben zum Firmensitz der Hersteller können aus dieser Tabelle entnommen werden und werden daher im Fließtext nicht explizit erwähnt.

2.1.3 Software, R-Packages und Online-Datenbanken

In diesem Abschnitt werden die in dieser Arbeit verwendeten Software-Anwendungen, R-Packages und Online-Datenbanken beschrieben. Die hier aufgeführte Software stand im Institut für Humangenetik der Christian-Albrechts-Universität zu Kiel zu Verfügung. Online-Datenbanken wurden u.a. für Assay-Design, Sequenz-Alignment und zum Erhalt von zusätzlichen Informationen wie z.B. Allelfrequenzen und genregulatorischen Elementen verwendet.

2.1.3.1 Software und R-Packages

In Tabelle 10 sind die in dieser Arbeit verwendeten Software-Anwendungen dargestellt.

Tabelle 10 Verwendete Software mit Angaben zum Hersteller und zu Anwendungsbereichen

Software (Version)	Hersteller	Anwendung
Affymetrix® Chromosome Analysis Suite (ChAS) (Version 2.1)	Affymetrix, Santa Clara, USA	Analyse von SNP-Arrays
BioDoc Analyzer	Applied Biosystems, Foster City, USA	Geldokumentation
Chromas Lite (Version 2.1.1)	Technelysium Pty Ltd, Brisbane, Australien	Darstellung von Elektropherogrammen der Sanger-Sequenzierung
Command Console™	Affymetrix, Santa Clara, USA	Gerätesoftware für die GeneChip Fluidics-Station 450
EpiTYPER™ Software (Version 1.2) mit Modulen: EpiTYPER Plate Editor und EpiTYPER Analyzer	Sequenom Inc., San Diego, USA	Plattenbelegung und Analyse der EpiTYPER-Technik
FastQC (Version 0.11.2)	Babraham Bioinformatics, Cambridge, England	Qualitätskontroll-Analyse von Rohdaten der RNA-Sequenzierung

Software (Version)	Hersteller	Anwendung
GenomeStudio™ Software (Version 2011.1) mit Methylation Analysis Modul (Version 1.9.0)	Illumina, San Diego, USA	Auswertung der 450K-Arrays
Lightcycler®480 SW (Version 1.5.1)	Roche, Rotkreuz, Schweiz	Analyse von qPCR-Daten
Microsoft Office 2007	Microsoft, Unterschleißheim, Germany	Ergebnisdokumentation und Auswertung
PyroMark® Assay Design (Version 2.0)	Qiagen, Hilden, Deutschland	Primerdesign für BSPS
PyroQ-CpG™ Software (Version 1.0.11)	Qiagen, Hilden, Deutschland	Auswertung der BSPS
Qlucore Omics Explorer (Version 3.0)	Qlucore, Lund, Schweden	Statistische Auswertung der 450K-Arrays
R-Software (Version 3.1.0)	R Development Core Team, Boston, USA	Statistische und graphische Auswertung
Sequencing Analysis Software™ (Version 6.0)	Applied Biosystems, Foster City, USA	Analyse von Sequenzen der Sanger-Sequenzierung

450K-Array: Infinium HumanMethylation450 BeadChip Array; SNP-Array: Genome-Wide Human SNP Array 6.0; qPCR: quantitative Real-Time PCR; BSPS: Bisulfit-Pyrosequenzierung; RNA-Sequenzierung: Gesamt-Transkriptom-Sequenzierung. In dieser Arbeit erfolgt lediglich bei der ersten Erwähnung dieser Software im Fließtext eine Angabe zum Hersteller und ggf. zur Software-Version. Angaben zum Hersteller können aus dieser Tabelle entnommen werden und werden daher im Fließtext nicht explizit erwähnt.

In Tabelle 11 sind die in dieser Arbeit mittels R-Software verwendeten R-Software-Anwendungen und R-Packages aufgelistet.

Tabelle 11 R-Software-Anwendungen und R-Packages mit URL-Angabe, Zitierung und Referenz

Name	URL	Zitierung	Referenz
Bioconductor (BiocLite)	https://bioconductor.org/biocLite	Huber, W., Carey, V.J., Gentleman, R., Anders, S., Carlson, M., Carvalho, B.S., Bravo, H.C., Davis, S., Gatto, L., Girke, T., et al. (2015). Orchestrating high-throughput genomic analysis with Bioconductor. https://bioconductor.org .	(Huber et al., 2015) [252]
Biostrings	https://bioconductor.org/packages/release/bioc/html/Biostrings.html	H. Pages, P. Aboyoun, R. Gentleman and S. DebRoy (). Biostrings: String objects representing biological sequences, and matching algorithms. R package version 2.32.1.	(Pages et al.) [253]
BSgenome	http://bioconductor.org/packages/release/bioc/html/BSgenome.html	Herve Pages (). BSgenome: Infrastructure for Biostrings-based genome data packages. R package version 1.32.0.	(Pages et al.) [253]
EdgeR	https://bioconductor.org/packages/release/bioc/html/edgeR.html	Robinson MD, McCarthy DJ and Smyth GK (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data.	(Robinson et al., 2010) [254]
MassArray	http://www.bioconductor.org/packages/release/bioc/html/MassArray.html	Reid F. Thompson and John M. Greally (2009). MassArray: Analytical Tools for MassArray Data. R package version 1.16.0.	(Thompson et al., 2009) [255]
RnBeads	http://rnbeads.mpi-inf.mpg.de/	Yassen Assenov, Fabian Mueller, Pavlo Lutsik, Joern Walter, Thomas Lengauer and Christoph Bock (2014). Comprehensive Analysis of DNA Methylation Data with RnBeads, http://rnbeads.mpi-inf.mpg.de .	(Assenov et al., 2014) [256]

In dieser Arbeit erfolgt lediglich bei der ersten Erwähnung dieser Software im Fließtext die Angabe der Referenz. Weitere Angaben zu URL und Zitierung können aus dieser Tabelle entnommen werden und werden daher im Fließtext nicht explizit erwähnt.

2.1.3.2 Online-Datenbanken und Online-Software-Anwendungen

Die in der Arbeit verwendeten Online-Datenbanken und Online-Software-Anwendungen sind in Tabelle 12 zusammengefasst.

Tabelle 12 Online-Datenbanken und Online-Software-Anwendungen mit URL-Angabe und ggf. Referenz

Name	URL	Referenz
1000 Genomes Project (1000G)	www.1000genomes.org/	(Consortium, 2010) [257]
BLAT (UCSC BLAT-Analyse)	http://genome.ucsc.edu/cgi-bin/hgBlat?hgsid=446995209_nAH1Ca4baTk7yZiPrwgsxtApAcHB&command=st art	(Kent, 2002) [258]
dbSNP (Short Genetic Variations)	http://www.ncbi.nlm.nih.gov/SNP/	(Sherry et al., 2001) [259]
DGV (Database of Genomic Variants)	http://dgv.tcag.ca/dgv/app/home	(MacDonald et al., 2014) [260]
ELM (Eukaryotic Linear Motif resource, Functional Sites in Protein)	http://elm.eu.org/	(Dinkel et al., 2012) [261]
EMBOSS Needle (Pairwise Sequence Alignment)	http://www.ebi.ac.uk/Tools/psa/emboss_needle/	-
EpiDesigner (Beta)	http://www.epidesigner.com/	-
Exome Aggregation Consortium (ExAC) Version 0.3 (grundsätzlich auf europäische (nicht-finnische) Population bezogen)	http://exac.broadinstitute.org/	-
GeneCards®	http://www.genecards.org/	(Rebhan et al., 1998) [262]
GeneMANIA Gennetzwerk-Analyse-Tool	http://www.genemania.org	(Warde-Farley et al., 2010) [263]
Genic Intolerance (RVIS)	http://genic-intolerance.org/	(Petrovski et al., 2013) [264]
HUGO Gene Nomenclature Committee (HGNC)	http://www.genenames.org/	(Wain et al., 2002) [265]
Human Genome Variation Society (HGVS)	http://www.hgvs.org/mutnomen/	(den Dunnen and Antonarakis, 2000) [266]
Mouse Genome Informatics (MGI)	http://www.informatics.jax.org/	(Shaw, 2009) [267]
Mutation Taster	http://www.mutationtaster.org/	(Schwarz et al., 2010) [268]
NCBI (National Center for Biotechnology Information)	http://www.ncbi.nlm.nih.gov/	-
NCBI Conserved Domains Database (CDD)	http://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml	-
NCBI Gene (Gendatenbank)	http://www.ncbi.nlm.nih.gov/gene/	-
NCBI PubMed	http://www.ncbi.nlm.nih.gov/pubmed	-
NHLBI Exome Sequencing Project, Exonic Variant Server (EVS)	evs.gs.washington.edu/	-
OligoCalc	http://biotools.nubic.northwestern.edu/OligoCalc.html	(Kibbe, 2007) [269]
Online Mendelian Inheritance in Man (OMIM)	http://www.omim.org/	(Hamosh et al., 2005) [270]
Pathway Interaction Database (PID)	http://pid.nci.nih.gov/	(Schaefer et al., 2009) [271]

Name	URL	Referenz
Polyphen-2 (Polymorphism Phenotyping v2)	http://genetics.bwh.harvard.edu/pph2/	(Adzhubei et al., 2013) [272]
Reactome	http://www.reactome.org/	(Haw et al., 2011) [273]
STRING Protein-Protein Interaktions-Netzwerke	http://string-db.org/	(Franceschini et al., 2013) [274]
UCSC Genome Bioinformatics	http://genome.ucsc.edu/	(Kent et al., 2002) [251]
UCSC Genome Bioinformatics <i>Track</i> : Transcription Factor ChIP-seq (161 factors) from ENCODE with Factorbook Motifs	http://genome.ucsc.edu/ <i>Track</i> : Integrated Regulation from ENCODE/Txn Factor ChIP	(Gerstein et al., 2012) [275], (Wang et al., 2012) [276], (Wang et al., 2013a) [277]
Varbank Pipeline v.2.3	https://varbank.ccg.uni-koeln.de/	-
Variant Effect Predictor (VEP)	http://grch37.ensembl.org/Homo_sapiens/Tools/VEP/Edit?db=core;tl=cGH5XL4XjwOkilie-466240	-
WEB-based GENE SeT AnaLysis Toolkit (WebGestalt)	http://bioinfo.vanderbilt.edu/webgestalt/	(Wang et al., 2013b) [278]

In dieser Arbeit erfolgen lediglich bei der ersten Erwähnung dieser Online-Datenbanken bzw. Online-Software-Anwendungen im Fließtext die Benennung des vollständigen Names, ein Verweis zu diesem Abschnitt 2.1.3.2 sowie ggf. eine Angabe zur Referenz. Nachfolgende Erwähnungen dieser Online-Datenbanken bzw. Online-Software-Anwendungen im Fließtext erfolgen mit Angabe der Namens-Abkürzung. Angaben zu URL und Referenz können aus dieser Tabelle entnommen werden und werden daher im Fließtext nicht explizit erwähnt.

2.2 Methoden

In diesem Teil der Arbeit werden die angewandten Methoden der verschiedenen Analysen erläutert. Ein Großteil der Protokolle basiert auf Standardarbeitsanweisungen des Instituts für Humangenetik der Christian-Albrechts-Universität zu Kiel, wo molekulargenetische Untersuchungen nach Bestimmungen der DIN ES ISO 15189:2007/2013 durchgeführt werden. Die im folgenden Abschnitt aufgeführten Methoden wurden eigens durchgeführt, wenn nicht anders angegeben. Die Analysen wurden bei Raumtemperatur durchgeführt, sofern nicht anders beschrieben. Alle Analysen wurden unter Labor-sterilen Bedingungen durchgeführt.

2.2.1 Molekulargenetische Analysen von DNA

Im folgenden Abschnitt werden die Methoden zu den molekulargenetischen Analysen von DNA beschrieben. In Abbildung 7 ist dies schematisch als Übersicht dargestellt.

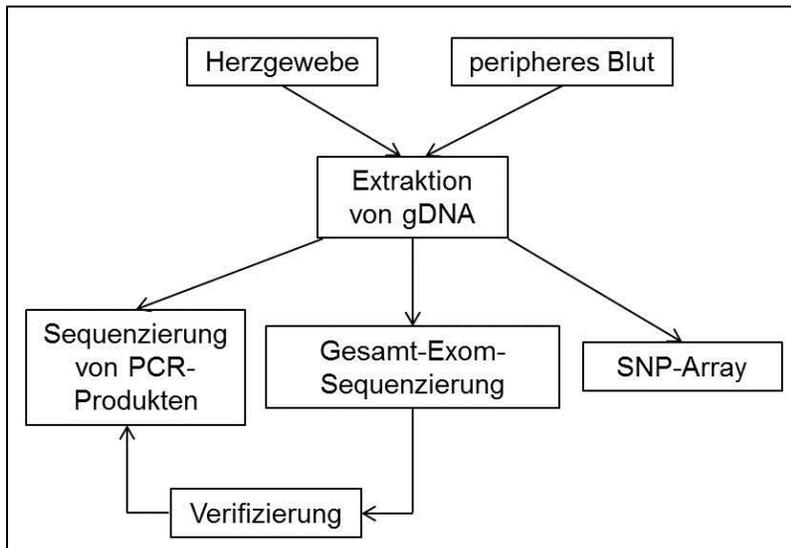


Abbildung 7 Schematische Darstellung der molekulargenetischen Analysen von DNA. gDNA: genomische DNA, SNP-Array: Genome-Wide Human SNP Array 6.0 (Affymetrix)

2.2.1.1 Extraktion von genomischer DNA

Genomische DNA (gDNA) wurde aus Herzgewebe und peripherem Blut (pB) extrahiert. Die dazu verwendeten Protokolle werden im Folgenden aufgeführt.

Zur **Extraktion von gDNA aus Herzgewebe** (ca. 5 – 15 mg) wurden Reagenzien des Gentra® Puregene® DNA Isolations-Kits (Qiagen) verwendet. Das Herzgewebe wurde zunächst mechanisch durch die Verwendung eines Skalpells zerkleinert und in ein 1,5 ml Reaktionsgefäß

überführt. Anschließend erfolgte die Extraktion von gDNA nach Angaben des Herstellers (Gentra® Puregene® Handbook, Qiagen (Angaben zur Extraktion aus 5 – 10 mg Gewebe)), was im Folgenden kurz erläutert wird. Nach Aufnahme in Qiagen Cell Lysis Solution wurde die Gewebesuspension mithilfe eines Vortexgerätes (REAX 2000, Heidolph) gemischt. Anschließend wurde Proteinase K hinzugegeben und bei 55 °C für 16 - 72 Stunden (h) in einem Thermoblock (UNO-Thermoblock, Biometra) inkubiert. Zu dem Lysat wurde dann RNase A hinzugegeben und 15 – 60 Minuten (min) bei 37 °C im Thermoblock (UNO-Thermoblock, Biometra) inkubiert. Die Proteinfällung erfolgte durch Zugabe von Qiagen Protein Precipitation Solution. Anschließend wurde die Suspension für ca. 1 min mittels Vortexgerät (REAX 2000, Heidolph) gemischt und dann auf Eis für mind. 10 min inkubiert. Nach anschließender Zentrifugation (4 °C) für 15 - 30 min bei 3515 g (Biofuge primoR, Heraeus) wurde der Überstand in ein neues 1,5 ml Reaktionsgefäß überführt. Die DNA-Präzipitation erfolgte durch Hinzufügen von 100 %igem Isopropanol (Sigma Aldrich) im 1:1 Volumenverhältnis. Die DNA konnte durch leichtes Schwenken des 1,5 ml Reaktionsgefäßes gefällt werden und der nun sichtbare weiße DNA-Faden wurde in ein neues 1,5 ml Reaktionsgefäß mit 70 %igem Ethanol (Details siehe 2.1.2.3) überführt, durch fünf-minütiges Zentrifugieren (Biofuge primoR, Heraeus) gewaschen und der Überstand verworfen. Das DNA-Pellet wurde bei Raumtemperatur getrocknet (ca. 2 h) und, je nach Größe des Pellets, in 20 - 100 µl ddH₂O aufgenommen und mind. 16 h gelöst.

Für die **Extraktion von gDNA aus pB** wurde das Gentra® Puregene® Blood Kit (Qiagen) verwendet. Im Wesentlichen entsprechen die Schritte denen der Extraktion von gDNA aus Gewebe und werden daher nur kurz beschrieben. Die Extraktion von gDNA erfolgte nach Angaben des Herstellers (Gentra® Puregene® Handbook, Qiagen (Angaben zur Extraktion aus 3 ml pB)). Hierbei erfolgte zunächst die Lyse von Erythrozyten und Leukozyten durch Hinzufügen von Qiagen RBC Lysis Solution zu 3 ml einer pB-Probe. Nach 10-minütigem Schwenken wurde die Lösung für 10 min bei 391 g zentrifugiert (Megafuge 1.OR, Heraeus). Der Überstand wurde verworfen und das Pellet mehrfach mit Qiagen RBC Lysis Solution gewaschen, bis dieses eine weißliche Färbung zeigte. Nach Aufnahme in Qiagen Cell Lysis Solution wurde die Zellsuspension für 16 – 72 h bei 37 °C im Hybridisierungsöfen (Biometra) inkubiert. Anschließend erfolgten auch hier Proteinfällung, DNA-Präzipitation und Elution der DNA wie oben beschrieben mit den entsprechenden Mengen der Lösungen und Chemikalien nach Angaben des Herstellers zur Extraktion aus 3 ml pB (Gentra® Puregene® Handbook, Qiagen).

2.2.1.2 Quantifizierung von DNA

Die DNA-Konzentration wurde unter Anwendung des Qubit® dsDNA BR bzw. des Qubit® dsDNA HS Assays (Life Technologies) mittels Qubit® 2.0 Fluorometer (Life Technologies) gemessen. Die Quantifizierung der DNA erfolgt hierbei durch die Messung eines Fluoreszenzsignals, welches bei spezifischer Bindung eines fluoreszierenden Farbstoffs an doppelsträngige DNA (dsDNA) emittiert wird. Im Gegensatz zu konventionellen spektrophotometrischen Messungen können bei dieser Methode Effekte durch Kontaminationen, wie z.B. Restmengen von RNA, auf ein Minimum reduziert werden [279]. Die Messungen erfolgten entsprechend des Protokolls nach Herstellerangaben (Life Technologies).

2.2.1.3 Amplifikation von DNA-Fragmenten durch PCR

Die PCR-Amplifikation (*polymerase chain reaction* (Polymerase-Kettenreaktion)) von DNA-Fragmenten erfordert im ersten Schritt ein Primer-Design spezifisch für die zu analysierende DNA-Sequenz. Hierzu wurden 150 – 300 bp *upstream* und *downstream* des zu analysierenden Sequenzabschnitts Primersequenzen (für eine Amplifikatlänge von 300 - 600 bp) mit einer Primerlänge von 18 - 25 bp ausgewählt. Dies erfolgte anhand der gDNA-Sequenz, ermittelt über die UCSC Genome Bioinformatics Datenbank [251] (siehe 2.1.3.2). Mithilfe der Software OligoCalc [269] (siehe 2.1.3.2) wurde die prädiktive Schmelztemperatur (Richtwert: 60 – 65 °C) und der GC-Gehalt (Richtwert: 50 – 65 %) überprüft und ggf. alternative Primersequenzen getestet. Die Spezifität der zu analysierenden Sequenz wurde mithilfe der UCSC BLAT-Analyse (siehe 2.1.3.2) anhand der Genom-Referenz GRCh37/hg19 überprüft. Falls SNPs (*single nucleotide polymorphisms* (Einzel-Nukleotid-Polymorphismen)) in der Primer-Sequenz auftraten, wurde ein neues Primer-Design durchgeführt.

Die PCR-Amplifikation erfolgte durch die Verwendung des AmpliTaq Gold® Polymerase Systems (Life Technologies). Im Fall von GC-reichen DNA-Fragmenten wurde das AccuPrime™ Polymerase System (Life Technologies) verwendet. Im Folgenden werden die Reaktionsansätze aufgeführt.

Standardmäßig wurde ein 25 µl-Reaktionsmix in 0,2 ml Reaktionsgefäßen angesetzt:

AmpliTaq Gold® Polymerase System 25 µl-Reaktionsansatz		AccuPrime™ Polymerase System 25 µl-Reaktionsansatz	
1 µl	gDNA [100 ng/µl]	1 µl	gDNA [100 ng/µl]
0,5 µl	Vorwärts-Primer [20 pmol/µl]	0,5 µl	Vorwärts-Primer [20 pmol/µl]
0,5 µl	Rückwärts-Primer [20 pmol/µl]	0,5 µl	Rückwärts-Primer [20 pmol/µl]
1 µl	dNTP ¹⁴ -Lösung [5 mM/Nukleotid]	2,5 µl	10x Puffer
2,5 µl	10x Puffer II	0,3 µl	Accuprime Polymerase
2 µl	MgCl ₂ [25 mM]	20,2 µl	ddH ₂ O
0,2 µl	AmpliTaq Gold Polymerase		
17,3 µl	ddH ₂ O		

Je nach gDNA-Konzentration der Proben wurde die Menge an ddH₂O angepasst, sodass das finale Reaktionsvolumen von 25 µl gleich blieb.

Im Folgenden ist ein typisches PCR-Programm dargestellt:

Denaturierung	95 °C	5 min	} 36 Zyklen
	95 °C	30 s	
Primerhybridisierung	55 - 65 °C	30 s	
Elongation	68 °C AccuPrime™ Polymerase/ 72 °C AmpliTaq Gold® Polymerase	30 s	
	68 °C bzw. 72 °C	5 min	
Kühlen	4 °C	unendlich	

Um ein spezifisches PCR-Produkt zu erhalten, erfolgte die Primerhybridisierung bei möglichst hoher Temperatur (daher mit 55 – 65 °C angegeben). Die Hybridisierungstemperatur für die jeweiligen Primerpaare wurde in Test-Reaktionen mit anschließender Agarose-Gelelektrophorese ermittelt. Zusätzlich wurden ggf. zum Reaktionsansatz MgCl₂ (Life Technologies) oder Betain-Lösung (Sigma Aldrich) zur Optimierung der PCR-Reaktion hinzugegeben. Die PCR-Produkte wurden mittels Agarose-Gelelektrophorese (siehe 2.2.1.4) überprüft.

¹⁴ Desoxynukleosidtriphosphat

2.2.1.4 Agarose-Gelelektrophorese

Die Agarose-Gelelektrophorese diente der Überprüfung von PCR-Produkten hinsichtlich Spezifität, Quantität und Amplifikatgröße. Je nach erwarteter Größe des PCR-Produkts wurden unterschiedliche Agarose-Konzentrationen gewählt: ein 3 %-Gel bei DNA-Fragmenten von ca. 100 bp, ein 2 %-Gel bei DNA-Fragmenten < 1000 bp oder ein 1 %-Gel bei DNA-Fragmenten > 1000 bp. Die Agarose (Biozym) wurde mit 1 x TBE-Puffer (Details siehe 2.1.2.3) durch Aufkochen in der Mikrowelle gelöst, auf Handwärme abgekühlt und mit Ethidiumbromid (Sigma Aldrich) für eine finale Ethidiumbromidkonzentration von 1 µM versehen. Das Agarose-Ethidiumbromid-Gemisch wurde anschließend in eine Gelkammer (peqLab, peqLab Biotechnologie) mit entsprechendem Taschenkamm überführt. Nach Verfestigung des Gels wurden 10 µl der zu analysierenden DNA-Fragmente mit 3,5 µl Agarose-Gel-Ladepuffer (Details siehe 2.1.2.3) versehen und das hergestellte Agarose-Gel beladen, welches sich in 1 x TBE-Puffer in der Agarose-Gelelektrophorese-Vorrichtung befand. Als Längenstandard wurde bei 1 %-Gelen der HyperLadder™ 1kb (Bioline) und bei 2 %- bzw. 3 %-Gelen der TriDye™ 100 bp DNA Ladder (New England Biolabs) verwendet. Bei 1 %-Gelen wurde eine Spannung von 5 V/cm und bei 2 %- bzw. 3 %-Gelen eine Spannung von 10 V/cm angelegt, bis die Längenstandard-Banden deutlich getrennt sichtbar waren. Die Dokumentation der Gele erfolgte mittels BioDoc Analyzer (Biometra). Die DNA-Fragmentgrößen konnten durch Vergleich zum Längenstandard ermittelt werden.

2.2.1.5 Aufreinigung von PCR-Produkten

PCR-Produkte wurden zur Vorbereitung auf die PCR-basierte DNA-Sequenzierung mittels MinElute® 96 UF PCR Purification Kit (Qiagen) zum Entfernen von restlichen Nukleotiden, Polymerase und Salzen aufgereinigt. Hierzu wurden 40 µl ddH₂O zum PCR-Produkt (ca. 15 µl Restmenge nach der Überprüfung mittels Agarose-Gelelektrophorese) hinzugefügt und dieses Gemisch in eine 96-Well Platte des MinElute® 96 UF PCR Purification Kits (Qiagen) überführt. Mithilfe einer Vakuumpumpe (Druck: 23 – 25 mm Hg, Millivac™ Vacuum Pump, Millipore) wurde die Flüssigkeit durch die Membran der 96-Well Platte abgesaugt, bis diese komplett trocken war. Durch Hinzufügen von 50 µl ddH₂O zu jedem Well und erneutes Anlegen eines Vakuums (s.o.) wurden die PCR-Produkte gewaschen. Anschließend wurde 20 µl ddH₂O zu jedem Well hinzugegeben, die 96-Well Platte mit einer Plastikfolie versiegelt und auf einem Plattenschüttler (IKA® MS3 basic, IKA) für 5 min bei 1000 U/min gelöst. Die aufgereinigten PCR-Produkte wurden dann in ein neues 0,2 ml Reaktionsgefäß überführt.

2.2.1.6 Sequenzierung von PCR-Produkten

Die Sequenzierung von PCR-Produkten, die zur Verifizierung von Ergebnissen aus der Gesamt-Exom-Sequenzierung (2.2.1.7) durchgeführt wurde, erfolgte nach der durch Sanger et al. (1977) [280] beschriebenen Didesoxymethode-Methode (im Folgenden Sanger-Sequenzierung abgekürzt). Die Didesoxymethode-Methode, auch Kettenabbruch-Synthese bezeichnet, basiert auf dem Einbau von Fluoreszenz-Farbstoff-markierten Didesoxynukleosidtriphosphaten (ddNTPs), welche durch das Fehlen der 3'-Hydroxygruppe eine Strangverlängerung durch die DNA-Polymerase verhindern [281]. Die entstehenden Kettenabbruchprodukte werden mittels Kapillarelektrophorese aufgetrennt und die unterschiedlichen Fluoreszenzfarben der vier ddNTPs detektiert, sodass ein Elektropherogramm der jeweiligen DNA-Sequenzierung erstellt wird. Die Sequenzier-PCR-Reaktion wurde mittels Big Dye Terminator® v1.1 Cycle Sequencing Kit (Applied Biosystems) durchgeführt.

Standardmäßig wurde ein 20 µl-Reaktionsmix in 0,2 ml Reaktionsgefäßen angesetzt:

Big Dye Terminator® v1.1 Cycle Sequencing Kit	
20 µl-Reaktionsansatz	
1 µl	Big Dye Terminator
2 µl	5x Big Dye Terminator Sequenzierpuffer
1 µl	Primer [20 pmol/µl]
1 µl	aufgereinigtes PCR-Produkt
15 µl	ddH ₂ O

Dieser Reaktionsansatz wurde sowohl für den Vorwärts- als auch für den Rückwärts-Primer angesetzt, um eine Vorwärts- und eine Rückwärts-Sequenz zu erhalten. Die Menge an PCR-Produkt, die für die Sequenzier-PCR-Reaktion eingesetzt wurde, variierte je nach Bandenintensität auf dem Agarose-Gel zwischen 1 µl – 3 µl. Die ddH₂O-Menge wurde entsprechend des 20 µl Reaktionsvolumens angepasst.

Die PCR-Reaktion erfolgte nach folgendem Programm:

Denaturierung	96 °C	2 min	} 25 Zyklen
	96 °C	10 s	
Primerhybridisierung	57 °C	5 s	
Elongation	60 °C	4 min	
Kühlen	4 °C	unendlich	

Die Sequenzier-PCR-Reaktion wurde anschließend mittels Montage™ SEQ96 Sequencing Reaction Cleanup Kit (Merck Millipore) zum Entfernen von restlichen Nukleotiden, Puffer und Salzen aufgereinigt. Hierzu wurden 20 µl Montage Injection Solution (Merck Millipore) zum PCR-Produkt hinzugefügt und dieses Gemisch in die Montage 96-Well Platte (Merck Millipore) überführt. Mit Hilfe einer Vakuumpumpe (Druck: 23 – 25 mm Hg, Millivac™ Vacuum Pump, Merck Millipore) wurde die Flüssigkeit durch die Membran der 96-Well Platte abgesaugt bis diese komplett trocken war. Durch Hinzufügen von 20 µl Montage Injection Solution (Merck Millipore) zu jedem Well und erneutes Anlegen eines Vakuums (s.o.) wurden die PCR-Produkte gewaschen. Anschließend wurde 20 µl ddH₂O zu jedem Well hinzugegeben, die 96-Well Platte mit einer Plastikfolie versiegelt und auf einem Plattenschüttler (IKA® MS3 basic, IKA) für 5 min bei 1000 U/min gelöst. Die aufgereinigten PCR-Produkte wurden dann in ein neues 0,2 ml Reaktionsgefäß überführt.

Die Sequenzierung erfolgte am ABI PRISM® 3130 Genetic Analyzer (Life Technologies). Hierzu wurde die aufgereinigte Sequenzier-PCR-Reaktion 4x – 10x verdünnt, abhängig von der Bandenintensität der vorherigen PCR-Reaktion im Agarosegel. Über die Sequencing Analysis Software™ (Version 6.0, Applied Biosystems) konnte die Injektionszeit von 5 s bis 15 s, je nach Sequenzierproduktkonzentration, und die Laufzeit von 18 min bis 30 min, je nach DNA-Fragmentlänge, vorgegeben werden. Die Analyse der DNA-Sequenzen in Form von Elektropherogrammen erfolgte anschließend mittels Sequencing Analysis Software™ (Applied Biosystems). Zur Qualitätskontrolle wurden die DNA-Sequenzen mit Referenzsequenzen (GRCh37/hg19) der Online-Datenbank NCBI (siehe 2.1.3.2) abgeglichen. Mutationen wurden nach HGVS-Nomenklatur (Human Genome Variation Society [266], siehe 2.1.3.2) beschrieben.

2.2.1.7 Gesamt-Exom-Sequenzierung

Die Gesamt-Exom-Sequenzierung (*Whole Exome Sequencing* (WES)) dient der Sequenzierung aller bekannter protein-kodierenden Gene eines Genoms. Am Cologne Center for Genomics (CCG) erfolgte durch wissenschaftliche Mitarbeiter die WES-Analyse von jeweils 1 µg DNA pro Probe. Die generellen Abläufe einer WES bestehen aus: (1) *Library*-Präparation zur Anreicherung exonischer Sequenzen der DNA-Proben, (2) Sequenzierung, hier mittels Illumina HiSeq 2000 System und (3) Computer-basierte Methoden, bestehend aus dem Alignment zur humanen Genom-Referenz, dem *Variante-Calling* und der Annotation zu SNP-Datenbanken. Im Folgenden wird die Methodik kurz erläutert.

Im Anschluss an die Fragmentierung der gDNA (hier mittels Sonifizierung) erfolgt die **Library-Präparation**, die in dieser Arbeit mittels SeqCap EZ Human Exome Library v2.0 Enrichment-Kit (Roche NimbleGen) durchgeführt wurde, welches die Analyse von mehr als 20.000 Genen, sowie microRNA-Genen ermöglicht. Um *blunt ends*, d.h. DNA-Fragment-Enden ohne 5'/3'-Überhänge, zu generieren, wird ein „*End Repair*“ mittels T4-DNA-Polymerase (3'-5'-Exonuklease- und 5'-3'-Polymerase-Aktivität) durchgeführt. Anschließend erfolgt die 5'-Phosphorylierung mittels T4-Polynukleotid-Kinase und *dA-Tailing* mittels Taq-Polymerase. Dies erleichtert die nachfolgende Ligation mit Sequenzier-Adaptoren, die die nötigen Elemente zur Phasen-Immobilisierung der späteren Sequenzierung enthalten. Im nächsten Schritt erfolgt die Hybridisierung an einen biotinylierten Oligo-Pool (DNA- oder RNA-*Baits*), der komplementär zu den exonischen Ziel-Sequenzen ist. Nach der Selektion der spezifischen Exom-DNA-Fragmente mittels Biotin-Streptavidin basiertem *Pulldown*-Assay erfolgt ein Waschvorgang zur Entfernung ungebundener Fragmente. Nach anschließender PCR-Amplifikation und Qualitätskontrolle kann die erstellte *Library* sequenziert werden.

In dieser Arbeit erfolgte die **Sequenzierung** mittels Illumina HiSeq 2000-System, basierend auf der sogenannten zyklisch-reversiblen Terminations-Methode (CRT) [282]. Nach Denaturierung erfolgt die Bindung der DNA-Fragmente an eine sogenannte „*Flow Cell*“, die kovalent gebundene Vorwärts- und Rückwärts-Primer in hoher Dichte enthält. Letztgenannte Primer sind zu Enden eines Universal- und eines Indexadapters komplementär. Der Universaladapter dient der späteren Hybridisierung des Sequenzierprimers für die Sequenzierung, während der Indexadapter der Probenidentifizierung dient. Durch Hinzufügen von Nukleotiden (unmarkiert) und Enzymen erfolgt eine klonale „*Bridge Amplification*“ des immobilisierten DNA-*Templates* mit benachbarten Primern, sodass sich 100 - 200 Millionen Cluster von bis zu 1000 dsDNA-Kopien eines *Templates* bilden [282]. Nach anschließender Denaturierung erfolgt die Hybridisierung von Sequenzier-Primern an die Adapter-Sequenzen (*paired-end* Sequenzierung), sowie die Zugabe von reversibel markierten Nukleotiden und der DNA-Polymerase (CRT-Methode). Nach Inkorporation des fluoreszenz-markierten Nukleotids kommt es zum Kettenabbruch, indem eine chemische 3'-Blockierungsgruppe am Nukleotid eine weitere Synthese verhindert. Nicht-gebundene Nukleotide werden herausgewaschen und anschließend nach der Laseranregung die entsprechend emittierte Fluoreszenz (unterschiedlich pro Nukleotid) ein jedes Clusters detektiert. Fluorophore und Blockierungsgruppe werden abgespalten und durch Waschen entfernt. Dieser Sequenzierungs-Zyklus wird für jede nachfolgende Base des gesamten Fragments wiederholt. [282]

Im Folgenden werden die **Computer-basierten Methoden** zur weiteren Prozessierung der Sequenzier-Daten erläutert:

Entsprechend der Varbank-Pipeline v.2.3 erfolgte durch Bioinformatiker des CCG eine primäre Qualitätskontrolle (siehe <https://varbank.ccg.uni-koeln.de/> (2.1.3.2)), das Alignment zur humanen Genom-Referenz (GRCh37/hg19), das Identifizieren der Varianten (Varianten-Calling) und die Annotation der Varianten zu SNP-Datenbanken sowie zu Algorithmen der Mutationskonsequenz-Vorhersage. Zur Gesamt-Analyse der neun IAS-Exome wurde wie folgt gefiltert (Abbildung 8):

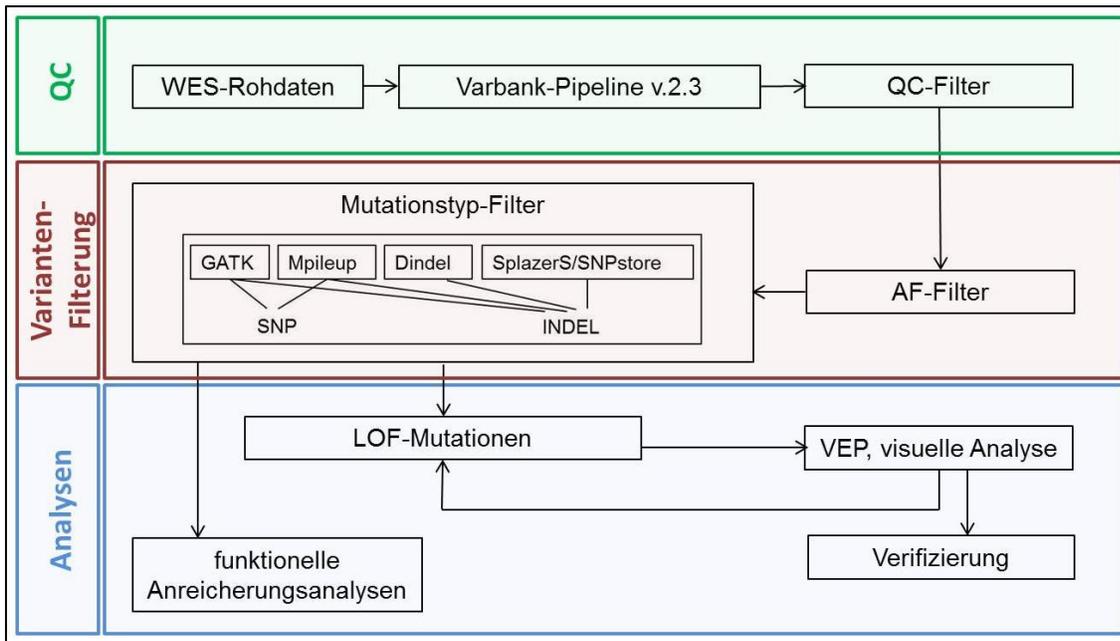


Abbildung 8 Filter-Pipeline und Daten-Prozessierung der WES-Analyse vom IAS-Gewebe von neun HLHS-Patienten. In der Abbildung sind die drei Ebenen der Datenauswertung – Qualitätskontrolle (QC), Varianten-Filterung und Analysen – dargestellt. Varbank-Pipeline v.2.3 beinhaltet: primäre QC, Alignment und Varianten-Calling; QC-Filter: Qualitätskontroll-Filter; AF-Filter: Filter nach Allelfrequenz; Mutationstyp-Filter: Filter nach Varianten, die die Proteinstruktur beeinflussen mit *Calling*-Software-spezifischer Selektion von SNPs und INDELS (*Calling* Algorithmen: GATK, Mpileup, Dindel, SplazerS/SNPstore); LOF: *Loss-Of-Function* (Frameshift-, Stop-, Start-Loss- und Splice-Site); VEP: Variant Effect Predictor (1000G-rel77); Visuelle Analyse: Visuelle Analyse der Reads des WES-Datensatzes.

Folgende Parameter der **Qualitätsfilterung (QC-Filter)** wurden in dieser Arbeit definiert (Abbildung 8):

- (1) minimale *Reads Coverage* von 6 *Reads*
- (2) Ausschluss von Genvarianten, die von weniger als zwei *Calling*-Software-Tools identifiziert wurden
- (3) Ausschluss von InHouseDB-Varianten (Epilepsy InhouseDB und Structural InhouseDB), die Pipeline-spezifische Artefakte darstellten

Zur stringenten Filterung seltener Genvarianten wurde in dieser Arbeit eine Allelfrequenz (AF) $\leq 0,01$ % als Filterkriterium des **AF-Filters** festgelegt. Dies basierte auf einer kombinierten AF der Datenbanken (siehe 2.1.3.2) dbSNP (build135) [259], 1000 Genomes Project (1000G, build20110521 [257]), Exonic Variant Server (EVS (NHLBI Exome Sequencing Project Version ESP6500)) und ExAC¹⁵ (Exome Aggregation Consortium, Version 0.3). Anschließend wurde nach **Mutations-Typ** (Abbildung 8) zur Selektion proteinstruktur-verändernder (funktioneller) Mutationen gefiltert, welcher folgende Parameter umfasste:

- (1) Ausschluss von Intron-Varianten, mit Ausnahme von Splice-Site-Mutationen (Position -1/+1 zu Exon)
- (2) Ausschluss synonymen Genvarianten
- (3) *Calling*-Software-spezifische Selektion von SNPs und INDELS (Kombination aus Insertion und Deletion):
 - a. nur SNPs, die mindestens durch die *Calling*-Software GATK und Mpileup identifiziert wurden
 - b. nur INDELS, die mindestens durch die *Calling*-Software GATK, Mpileup, Dindel und SplazerS/SNPstore identifiziert wurden

Anhand dieser gefilterten Liste wurden **Anreicherungsanalysen** (Details siehe 2.2.5.3) durchgeführt. Die zuvor nach Varbank-Pipeline v.2.3 annotierten LOF-Mutationen (*Loss-Of-Function*), d.h. Frameshift-, Stop-, Start-Loss- und Splice-Site-Mutationen, wurden mit Annotationen der **VEP-Datenbank** (Variant Effect Predictor (Ensembl, 1000G-rel77), siehe 2.1.3.2) abgeglichen. Eine **visuelle Analyse** beinhaltete eine Analyse der Güte der *Reads Coverage*, d.h. Richtung der *Read*-Sequenzen und Position der Mutation innerhalb der *Reads*, analysiert (Abbildung 8). Mittels Sanger-Sequenzierung (siehe 2.2.1.6) erfolgten die **Verifizierungen** von Mutationen.

Filtermethoden zur kombinierten Keimbahn- und somatischen Analyse

Zusätzlich zu den mittels WES analysierten IAS-Gewebeproben der neun HLHS-Patienten standen von einer anderen WES-Plattform Daten von pB-Proben eines Eltern-Index-Trios des zuvor an IAS-Gewebe analysierten HLHS-Falls HLHS_4 zur Verfügung (Tabelle 13, Familie 4). Hierbei wurde das Agilent SureSelect Human All Exon 50MB Kit (Agilent Technologies) verwendet. Die oben genannte Familie wurde durch Dr. med Marc-Phillip Hitz am Wellcome

¹⁵ In dieser Arbeit bezieht sich die ExAC-AF grundsätzlich auf ethnisch-gleiche Individuen: europäische (nicht-finnische) Population

Trust Sanger Institute (Cambridge, England) analysiert. Anhand dieser Familie erfolgte eine Analyse von IAS-spezifischen (somatischen) Mutationen (siehe 3.2.1.2). Die Prozessierung und Filterung der WES-Daten dieser kombinierten Keimbahn- und somatischen Analyse wird im Folgenden erläutert (Abbildung 9). Mittels Varbank-Pipeline v.2.3 und anschließender QC-Filterung erfolgte, wie zuvor beschrieben, die Prozessierung der WES-Rohdaten der IAS-Probe des HLHS-Falls HLHS_4 und der pB-Proben sowie des entsprechenden Eltern-Kind-Trios¹⁶ (Familie_4, siehe 2.2.1.7 (Tabelle 13)). Die AF-Filterung erfolgte mit einer $AF \leq 1\%$ (kombinierte AF, s.o.). Folgende Parameter des **IAS-Filters** (Abbildung 9) wurden definiert:

- (1) Selektion von Genvarianten, die ausschließlich bei der IAS-Gewebeprobe des Patienten, nicht jedoch in der pB-Probe des Patienten und dessen Eltern vorlagen
- (2) Ausschluss von Intron-Varianten
- (3) *Calling*-Software-spezifische Selektion von SNPs und INDELs (s.o.)

Anschließend erfolgte ein Abgleich der Varbank-Pipeline v.2.3-annotierten Mutationskonsequenzen mit Annotationen der VEP-Datenbank, sowie eine visuelle Analyse der Güte der *Reads Coverage* (s.o.). Die resultierenden IAS-spezifischen Mutationen wurden mittels Sanger-Sequenzierung (siehe 2.2.1.6) verifiziert.

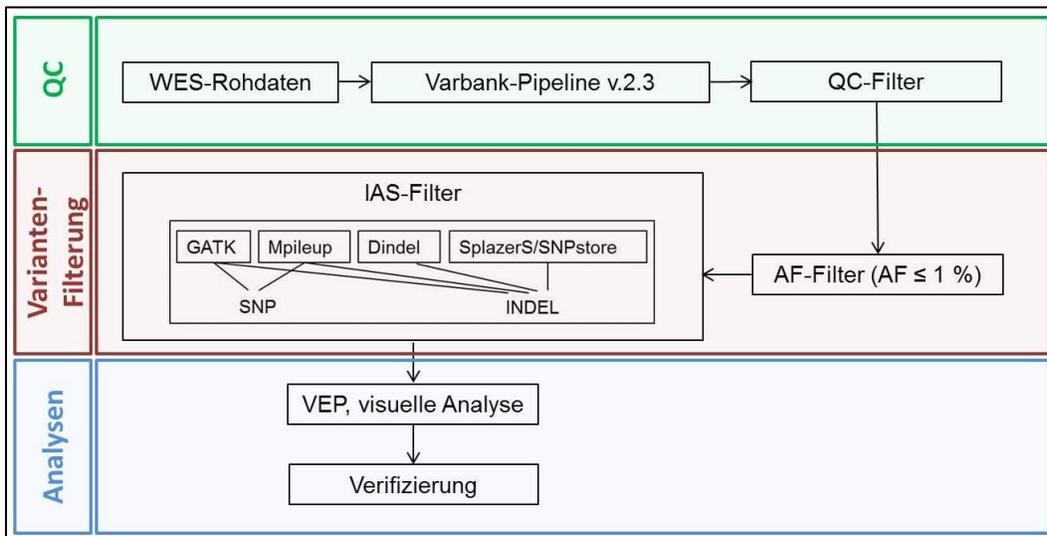


Abbildung 9 Filter-Pipeline und Daten-Prozessierung der Analyse IAS-spezifischer Mutationen der Familie 4 mit dem HLHS-Fall HLHS_4. In der Abbildung sind die drei Ebenen der Datenauswertung – Qualitätskontrolle (QC, entspricht Abbildung 8), Varianten-Filterung und Analysen – dargestellt. Varbank-Pipeline v.2.3 beinhaltet: primäre QC, Alignment und Varianten-*Calling*; QC-Filter: Qualitätskontroll-Filter; AF Filter: Filter nach Allelfrequenz (hier $AF \leq 1\%$); IAS-Filter: Filter für IAS-spezifische (nicht in pB vorkommende) Mutationen mit *Calling*-Software-spezifischer Selektion von SNPs und INDELs (*Calling* Algorithmen: GATK, Mpileup, Dindel, SplazerS/SNPstore); VEP: Variant Effect Predictor (1000G-rel77); Visuelle Analyse: Visuelle Analyse der *Reads* des WES-Datensatzes.

¹⁶ WES-Rohdaten stammen von einer anderen Sequenzierungsplattform mittels Agilent SureSelect Human All Exon 50MB Kit (siehe 2.2.1.7 (Tabelle 13))

Proben und Sequenzierungsplattformen

Weitere WES-Daten von pB-Proben einer Familie mit zwei HLHS-Fällen wurden durch den Kollegen Dr. med Marc-Phillip Hitz zur Verfügung gestellt (Tabelle 13, Familie 5). In dieser Arbeit erfolgten hierzu Analysen im Rahmen einer WES-Studie zu Familien mit linksventrikulären Ausflusstrakt-Obstruktionen (LVOTOs) (Hitz und Hoff et al., Publikation in Vorbereitung (siehe 11); Ergebnisse siehe 3.1.2.2). Eine Zusammenfassung der mittels WES analysierten Proben und Sequenzierungsplattformen ist in Tabelle 13 aufgeführt.

Tabelle 13 Auflistung der Proben und WES-Plattformen

Proben-ID	Familie	Gewebe-typ	Enrichment-Kit	Plattform	Durchführung Institut
D12/579 D12/581	1	pB	SeqCap EZ Human Exome Library v2.0 Kit <i>DNA-Bait</i> (Roche NimbleGen, Madison, USA)	Illumina HiSeq 2000	CCG Cologne Center for Genomics (Köln, Deutschland)
D12/775 D12/776	2				
D11/527 D11/528 D11/1087	3				
HLHS_12 HLHS_21 HLHS_22 HLHS_132 HLHS_4 (IAS-Probe zu HLHS_4_Index) HLHS_8 HLHS_13 HLHS_15 HLHS_35	-	IAS			
HLHS_4_Index (pB-Probe zu HLHS_4) HLHS_4_Mutter HLHS_4_Vater	4	pB	Agilent SureSelect Human All Exon 50MB Kit <i>RNA-Bait</i> (Agilent Technologies, Santa Clara, USA)	Illumina HiSeq 2000	Wellcome Trust Sanger Institute (Cambridge, England)
HLHS_130_Index-1 (pB-Probe zu HLHS_130) HLHS_130_Index-2 HLHS_130_Vater	5				

Familie: Familie siehe Abschnitt 2.1.1.1, Tabelle 2; Enrichment-Kit: Kit für die Anreicherung exonischer Sequenzen; pB: peripheres Blut; IAS: interatriales Septum; Plattform: Sequenzierungsplattform; in grau schattiert: diese Analysen wurden durch Dr. med Marc-Phillip Hitz am Wellcome Trust Sanger Institute (Cambridge, England) durchgeführt und die Daten wurden zur Verfügung gestellt.

Abgleich zu bekannten Genen, die in Assoziation mit Herzfehlbildungen identifiziert wurden

Kardiale Genlisten dienten dem Abgleich der in dieser Arbeit identifizierten Genen aus der WES-Analyse mit aus der Literatur bekannten Genen, die in Zusammenhang mit kardialen Anomalien gefunden wurden. Diese Genlisten wurden durch den Kollegen Dr. med. Marc-Phillip Hitz [283] im Vorfeld erstellt. Die Genlisten basieren auf Literaturrecherchen mittels NCBI PubMed-

Datenbank (siehe 2.1.3.2), sowie auf Listen, die mithilfe der MGI-Datenbank (Mouse Genome Informatics [267], siehe 2.1.3.2) generiert wurden. Entsprechend kardiovaskulärer HPO-Terms (Human Phenotype Ontology, siehe 2.1.3.2) wurden Gene mit bzw. ohne kardiovaskulärem Phänotyp im Mausmodell selektiert. Die Listen bestanden aus je einer Liste von Genen (1) ohne kardiovaskulärem Phänotyp im Mausmodell (N = 1351) als Kontrolldatensatz, (2) mit kardiovaskulärem Phänotyp im Mausmodell (N = 463) und (3) kardiovaskulärem humanen Phänotyp (N = 182) [283]. Die beiden letztgenannten Genlisten mit bekanntem kardialem Phänotyp werden im Folgenden als Herz-Genliste bezeichnet.

RVIS-Scoring-System

Zur Untersuchung des WES-Datensatzes wurde das **RVIS-Scoring-System** (RVIS: *Residual Variation Intolerance Score*; siehe 2.1.3.2) nach Petrovski et al. (2013) [264] als Analyseparameter verwendet. Der RVIS-Wert definiert pro Gen die Häufigkeit des Auftretens funktioneller Varianten relativ zur genomweit erwarteten Häufigkeit, basierend auf AF-Daten (noch nicht publiziert) der ExAC-Datenbank ($AF \leq 0.05\%$). Der RVIS-Perzentil-Score (im Folgenden RPS abgekürzt) kann Werte zwischen 0 und 100 einnehmen mit 0 = geringe Mutationstoleranz und 100 = hohe Mutationstoleranz gegenüber funktionellen Varianten des jeweiligen Gens. Funktionelle Genvarianten in Genen mit niedrigem RPS sind mit höherer Wahrscheinlichkeit von krankheitsrelevanter Bedeutung als solche Genvarianten in Genen mit hohem RPS. In dieser Arbeit wurde ein $RPS \leq 25$ entsprechend der Studie von Petrovski et al. (2013) [264] als Grenzwert für Gene mit geringer Mutationstoleranz definiert.

Die in der vorliegenden Arbeit identifizierten Genvarianten wurden entsprechend folgender Kriterien kategorisiert:

Tabelle 14 Definierte Gen-Kategorien der identifizierten Genvarianten und jeweilige Beschreibung der Gen-Kategorie

Gen-Kategorie	Beschreibung
potenziell krankheitsrelevant	seltene bzw. unbekannte funktionelle Genvariante ¹ bei einem Fall mit CHD
wahrscheinlich krankheitsrelevant	seltene bzw. unbekannte LOF-Mutation ² , oder funktionelle Genvariante mit <i>compound</i> heterozygoter ³ , homozygoter oder <i>de novo</i> Ausprägung bei einem Fall mit CHD + kardialer Phänotyp des Gens im Modellorganismus (z.B. Maus)
kausal	seltene bzw. unbekannte funktionelle Genvarianten ¹ mit <i>compound</i> heterozygoter, homozygoter oder <i>de novo</i> Ausprägung bei mehr als einem (nicht verwandtem) Fall mit CHD ⁴ + Genvarianten, die als <i>possibly</i> oder <i>probably</i> damaging nach PPH ⁵ eingestuft sind

¹Genvariante mit Veränderung der kodierten Aminosäuren; ²LOF-Mutation: *Loss-Of-Function*-Mutation (Funktionsverlust); ³*compound* heterozygot: zwei unterschiedlich mutierte Allele desselben Gens; ⁴oder einer anderen angeborenen Fehlbildung; ⁵PPH: Polyphen-2 (Polymorphism Phenotyping v2 [272], siehe 2.1.3.2)

2.2.1.8 SNP-Array Analyse

Der Genome-Wide Human SNP Array 6.0 (Affymetrix, (im Folgenden als SNP-Array bezeichnet)), ermöglicht die genomweite Analyse von mehr als 906.600 SNPs und 946.000 Kopienzahlvariation (CNVs). Der SNP-Array enthält auf dem Array-Trägermaterial 25 bp-Oligonukleotid-Sonden, die das A- oder B-Allel des jeweiligen SNPs spezifisch binden. Die Durchführung erfolgte nach Herstellerangaben (Affymetrix Genome-Wide Human SNP Nsp/Sty 6.0 User Guide) und wurde teilweise durch Claudia Becher, Technische Assistentin des Instituts für Humangenetik der Christian-Albrechts-Universität zu Kiel, unterstützt. Im Folgenden werden die Vorgänge – von der Probenvorbereitung bis zum Scannen – kurz erläutert. Hierbei wurde das Genome-Wide Human SNP Nsp/Sty Assay Kit 6.0 (Affymetrix) verwendet.

Vor der DNA-Hybridisierung auf den Array wurde zunächst die Fragmentierung mithilfe der Restriktionsenzyme (RE) NspI- und StyI (New England Biolabs) durchgeführt. Anschließend erfolgte eine RE-spezifische Adapterligation unter Verwendung einer T4 DNA-Ligase (New England Biolabs), sowie eine PCR-Amplifikation mit Adapter-spezifischen Primern (TITANUM® Taq Polymerase Kit (Clontech Laboratories Inc., Mountain View, USA), welche mittels Agarose-

Gelelektrophorese (siehe 2.2.1.4) zur Qualitätskontrolle überprüft wurde. Die PCR-Produkte wurden dann mithilfe von magnetischen *Beads* (Agencourt® AMPure® XP Kit, Beckmann Coulter) aufgereinigt und enzymatisch mittels DNaseI (GeneChip® Fragmentation Reagent, Affymetrix) fragmentiert. Nach Überprüfung der Fragmentierung mittels Agarose-Gelelektrophorese (siehe 2.2.1.4) erfolgte die Biotin-Markierung (GeneChip® DNA Labeling Reagent und Terminal Deoxynucleotidyl Transferase, Affymetrix) und anschließend die Hybridisierung auf das Array-Trägermaterial, sodass die DNA-Fragmente an die entsprechenden Sonden gebunden wurden. Anschließend erfolgte mittels Waschpuffer (Details siehe 2.1.2.3) das Waschen zur Entfernung nicht-gebundener DNA-Fragmente und das Färben (mit Färbe-Puffer, Details siehe 2.1.2.3) des Arrays mithilfe der GeneChip® Fluidics Station 450 (Affymetrix), sodass die Fluoreszenzsignale dann mittels GeneChip® Scanner 3000 (Affymetrix) detektiert werden konnten.

Die Auswertung erfolgte anschließend mittels Affymetrix® Chromosome Analysis Suite Software Version 2.1 (ChAS, Affymetrix). Durch Vergleich der Genotypen von Indexpatient und Eltern konnte so hinsichtlich differenzieller Kopienzahlverluste- bzw. Zugewinnen oder auch hinsichtlich unbekannter Variationen analysiert werden. Hierbei wurden folgende Einstellungen gewählt: min. Markeranzahl = 20 und min. CNV-Länge = 100 kbp. In dieser Arbeit wurden CNVs, die in mehr als 20 Kontroll-Individuen nach DGV-Datenbank (Database of Genomic Variants [260], siehe 2.1.3.2) vorlagen und eine Überlappung von > 50 % mit bereits bekannten CNVs zeigten, als *benign* definiert.

2.2.2 DNA-Methylierungsanalysen

Im folgenden Abschnitt werden die verschiedenen Methoden der DNA-Methylierungsanalyse erläutert. Die aus unterschiedlichen Herzgewebeproben extrahierte gDNA wurde mithilfe dieser Techniken untersucht.

Im Anschluss an eine Natriumbisulfit-Behandlung (siehe Abschnitt 2.2.2.1) der zu untersuchenden gDNA-Proben, welche zu einer selektiven Basen-Konversion von methyliertem Cytosin zu Thymin führt [284, 285], erfolgte die Untersuchung der DNA-Methylierung der gDNA-Proben.

Die DNA-Methylierung wurde in dieser Arbeit durch drei verschiedene Methoden der DNA-Methylierungsanalyse quantifiziert:

- (1) Infinium HumanMethylation450 BeadChip Array (Illumina; 450K-Array abgekürzt), welcher eine genomweite DNA-Methylierungsanalyse von über 485.000 Loci¹⁷ ermöglicht
- (2) Bisulfit-Pyrosequenzierung (BSPS) zur PCR-basierten DNA-Methylierungsanalyse
- (3) MassARRAY-EpiTYPER-Technik (MassARRAY® EpiTYPER®, Sequenom Inc; EpiTyper-Technik abgekürzt), die mittels massenspektrometrischen Verfahrens PCR-basierte DNA-Methylierungsanalysen ermöglicht

Die beiden letztgenannten Methoden wurden zur Verifizierung von 450K-Array-Ergebnissen genutzt. In der vorliegenden Arbeit wurde die EpiTyper-Technik für die Anwendung am Institut für Humangenetik der Christian-Albrechts-Universität zu Kiel neu etabliert. In Abbildung 10 sind die in der Arbeit angewandten Methoden der DNA-Methylierungsanalyse schematisch als Übersicht dargestellt.

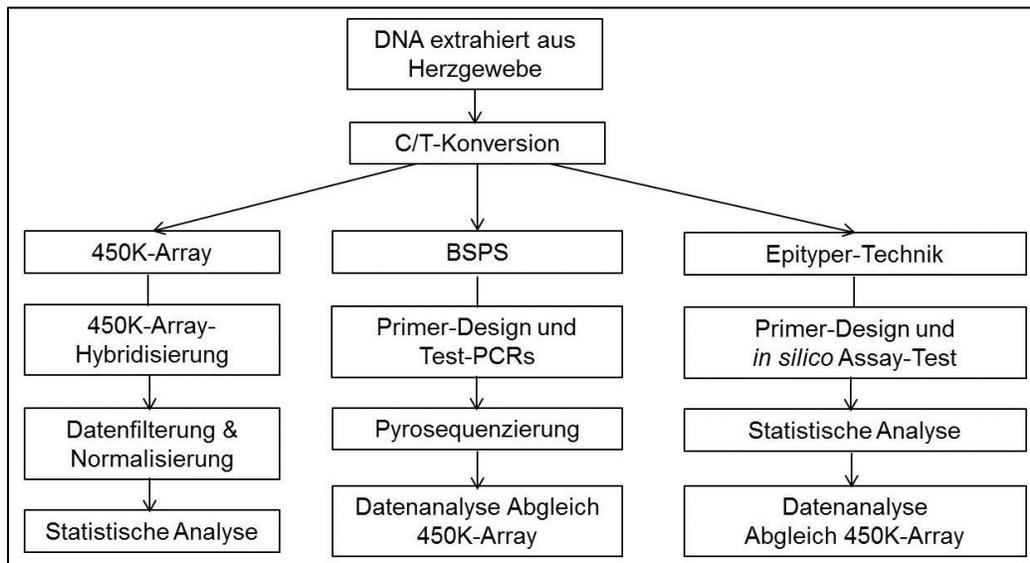


Abbildung 10 Schematische Darstellung der Methoden der DNA-Methylierungsanalyse. C/T-Konversion: Basen-Konversion von methyliertem Cytosin zu Thymin; 450K-Array: Infinium HumanMethylation450 BeadChip Array; BSPS: Bisulfit-Pyrosequenzierung; EpiTyper-Technik: MassARRAY-EpiTYPER-Technik.

¹⁷ Der Begriff Loci schließt Cytosin-Guanin-Dinukleotide (CpGs) sowie nicht-CpG-Loci (ch-Loci), die in humanen Stammzellen identifiziert wurden, als auch analysierbare SNP-Loci des Arrays ein.

2.2.2.1 Natriumbisulfit-Behandlung genomischer DNA

Die DNA-Methylierungsanalysen erforderten als ersten Schritt eine Natriumbisulfit-Behandlung der zu untersuchenden gDNA-Proben, um die Bestimmung der DNA-Methylierung durch die oben genannten Methoden zu ermöglichen. Die Methode basiert auf der Bisulfit-induzierten hydrolytischen Desaminierung von Cytosin mit anschließender alkalischer Desulfonierung, welche zu einer Konversion des Cytosins zu Uracil führt. Die Bisulfit-induzierte Konversion von 5-Methyl-Cytosin verläuft dagegen unter den gewählten Bedingungen deutlich langsamer [284, 285], sodass eine selektive Konversion von Cytosin, nicht jedoch von 5-Methyl-Cytosin zu Uracil stattfindet. Erfolgt im Anschluss an die Natriumbisulfit-Behandlung eine PCR-Amplifikation, so ersetzt die DNA-Polymerase die Uracil-Moleküle durch Thymin. Es wird dadurch ein C/T-Polymorphismus generiert. In einer anschließenden Analyse der DNA-Sequenz lässt sich somit unmethyliertes Cytosin als Thymin detektieren.

Die Natriumbisulfit-Behandlung erfolgte in dieser Arbeit mittels EZ DNA Methylation™ Kit (Zymo Research). Die Durchführung entsprach den Herstellerangaben (Instruction Manual EZ DNA Methylation™ Kit, Zymo Research) und wird im Folgenden kurz erläutert. Es wurden ausschließlich Puffer und Reagenzien des EZ DNA Methylation™ Kits (Zymo Research) verwendet.

Pro zu untersuchender DNA-Probe wurde 1 µg DNA in ein 0,2 ml Reaktionsgefäß überführt, auf ein Volumen von 45 µl mit ddH₂O verdünnt und 5 µl M-Dilution Puffer (Zymo Research) hinzugegeben. Anschließend erfolgte eine Inkubation bei 37 °C im Thermocycler (T3000 Thermocycler, Biometra). Nachdem zu jeder Probe 100 µl CT Conversion Reagenz (Zymo Research) hinzugegeben wurden, erfolgte die Konvertierungsreaktion (C zu U) unter folgenden Bedingungen:

95 °C	30 s	} 16 Zyklen
50 °C	60 min	
4 °C	unendlich	

Es wurden 400 µl M-Binding Puffer in Zymo-Spin™ IC Säulen in Auffanggefäßen vorgelegt, die konvertierte Probe hinzugegeben, gemischt, bei > 10.000 g für 30 s zentrifugiert (Mikro200, Hettich) und der Durchfluss verworfen. Anschließend wurden 100 µl M-Wash Puffer (Zymo Research) hinzugegeben und bei > 10.000 g für 30 s zentrifugiert (Mikro200, Hettich). Nachdem 200 µl M-Desulphonation Puffer (Zymo Research) hinzugegeben wurden, erfolgte eine

Inkubation bei Raumtemperatur für 15 - 20 min mit anschließender Zentrifugation (Mikro200, Hettich) bei > 10.000 g für 30 s. Durch Zugabe von 200 µl M-Wash Puffer (Zymo Research) und Zentrifugation (Mikro200, Hettich) bei > 10.000 g (30 s) wurde die Säule gewaschen. Dieser Vorgang wurde noch einmal wiederholt. Die Zymo-Spin™ IC Säule wurde dann in ein neues 1,5 ml Reaktionsgefäß überführt, 10 µl ddH₂O auf die Säulenmatrix gegeben und für 1 min inkubiert. Die Elution der konvertierten DNA erfolgte mittels Zentrifugation (Mikro200, Hettich) bei > 10.000 g (30 s). Die konvertierte DNA konnte dann für DNA-Methylierungsanalysen verwendet werden.

2.2.2.2 Array-basierte DNA-Methylierungsanalyse

Zur globalen DNA-Methylierungsanalyse wurde in dieser Arbeit der Infinium HumanMethylation450 BeadChip Array (Illumina; im Folgenden abgekürzt: 450K-Array) verwendet. Dieser 450K-Array ermöglicht die parallele genomweite DNA-Methylierungsanalyse von 485.577 verschiedenen Loci, die 21.231 Genen zuzuordnen sind. Diese 485.577 Loci sind unterteilt in: (1) 482.241 Cytosin-Guanin-Dinukleotid-Loci (CpG-Loci), (2) 3.091 nicht-CpG-Loci (in humanen Stammzellen (auch: ch-Loci) identifiziert) und ein methyliertes Cytosin in einem nicht-CpG-Kontext zeigen und (3) 65 CpG-Loci, die als SNPs vorliegen (auch: rs-Loci). Der Array basiert auf zwei verschiedenen *Bead*-Typ-Assays: dem „Infinium I Methylation Assay“ (Infinium I-Assay abgekürzt) und dem „Infinium II Methylation Assay“ (Infinium II-Assay abgekürzt) [286]. Der Infinium I-Assay basiert auf *Bead*-gebundenen Sonden, die an ihrem 3'-Ende eine DNA-Sequenz tragen, die entweder komplementär zum methylierten (Cytosin), oder zum unmethylierten (Thymin) CpG-Locus ist. Somit liegen hier zwei *Bead*-Typen vor: ein methylierter und ein unmethylierter. Die konvertierte¹⁸ DNA bindet an die komplementären *Bead*-Typen. Die Sonde wird im Zuge einer Einzelbasenelongation durch markierte Nukleotide (Adenin und Thymin: 2,4-Dinitrophenol (DNP) markiert; Cytosin und Guanin: Biotin markiert) hinter der CpG-Position verlängert, wenn eine Basenkomplementarität zur konvertierten DNA vorlag. Durch Bindung von Fluoreszenz-markierten Antikörpern erfolgt hier die Messung über einen gemeinsamen Farbkanal. Beim Infinium II-Assay liegt nur ein *Bead*-Typ pro CpG-Locus vor und die Sondensequenz endet hier auf dem Cytosin im CpG-Locus, sodass bei der Einzelbasenelongation entweder ein DNP-markiertes Adenin oder ein Biotin-markiertes Guanin inkorporiert wird. Als Fluoreszenzfarbstoffe dienen Cy3 (am Anti-Biotin Antikörper) und Cy5 (am

¹⁸ In der vorliegenden Arbeit ist die „Konversion“ oder „konvertierten Probe“ je nach Kontext definiert als eine C-zu-U-Konversion oder C-zu-T-Konversion (schließt Natriumbisulfit-Konversionsprozess (C zu U), sowie C zu T Austausch im Rahmen einer PCR-Amplifikation ein).

Anti-DNP-Antikörper), die im grünen (methyliert) bzw. roten (unmethyliert) Wellenlängenbereich nach Anregung emittieren. Aus den Signalintensitäten der gemessenen Fluoreszenz werden anschließend sogenannte AVG-Beta-Werte berechnet. Diese stellen in beiden *Bead*-Typ-Assays pro CpG-Locus das Verhältnis von methylierter Signalintensität zu der Summe von methylierter und unmethylierter Signalintensität dar (Werte von 0 (unmethyliert) bis 1 (vollständig methyliert)). Im Folgenden ist die Formel zur Bestimmung des AVG-Beta-Wertes aufgeführt:

$$AVG - Beta = \frac{\textit{Intensität methyliertes Signal}}{((\textit{Intensität methyliertes Signal} + \textit{unmethyliertes Signal}) + 100)}$$

Die Durchführung der Probenvorbereitung, bis zur Hybridisierung und Färbung des 450K-Arrays, erfolgte nach dem „Infinium HD Assay Methylation Protocol Guide: Infinium HD Methylation Manual Workflow“ (Illumina, Stand April 2011) und wird daher im Folgenden nur kurz erläutert. Ein Teil der Array-Hybridisierungen wurden durch Lorena Valles, Technische Assistentin des Instituts für Humangenetik der Christian-Albrechts-Universität zu Kiel, unterstützt. Hierzu wurde das Infinium HumanMethylation450 BeadChip Kit (Illumina) verwendet.

Es wurde 500 ng konvertierte DNA (siehe 2.2.2.1) pro zu analysierender Probe eingesetzt. Anschließend erfolgte eine Gesamt-Genom-Amplifikation (WGA), in der bei isothermalen Bedingungen (37 °C) mittels randomisierter Hexamerprimer und einer speziellen DNA-Polymerase das gesamte Genom amplifiziert wurde (20 – 24h). Hierbei findet der Austausch von U zu T durch die DNA-Polymerase statt, sodass ein C/T-Polymorphismus generiert wird. Die konvertierte DNA wurde enzymatisch fragmentiert, mittels Isopropanol (Sigma Aldrich) präzipitiert und in Hybridisierungspuffer (HumanMethylation450 BeadChip Kit, Illumina) resuspendiert. Nachdem die Proben auf den HumanMethylation450 BeadChip (Illumina) pipettiert wurden, erfolgte die Inkubation im Hybridisierungssofen (Illumina) für 16 - 22 h bei 48 °C. Hierbei findet eine Immobilisierung der DNA-Fragmente an komplementäre Sequenzen *Bead*-gebundener Sonden statt. Nicht-gebundene DNA-Fragmente wurden durch anschließende Waschschrte entfernt. Während der anschließenden Einzelbasenelongation wurden die Sonden-Sequenzen, bei vorliegender Basenkomplementarität der konvertierten DNA zur Sonde, durch die markierten Nukleotide (s.o.) verlängert. Nach Hinzufügen der Fluoreszenz-markierten Antikörper (s.o.) konnten die Fluoreszenzsignale mittels iScan (Illumina) detektiert und als IDAT-Dateien gespeichert werden.

2.2.2.2.1 Qualitätskontrolle und Auswertung

Mithilfe verschiedener Software-Anwendungen erfolgte die Auswertung der 450K-Array-Daten. Die Abläufe zu Qualitätskontrolle, Normalisierung, Filterung der Daten, sowie der statistischen Analyse sind in Abbildung 11 als Filter-Pipeline schematisch dargestellt. Diese Filter-Pipeline wurde für jede einzelne Analyse eines jeden Teilprojekts separat durchgeführt.

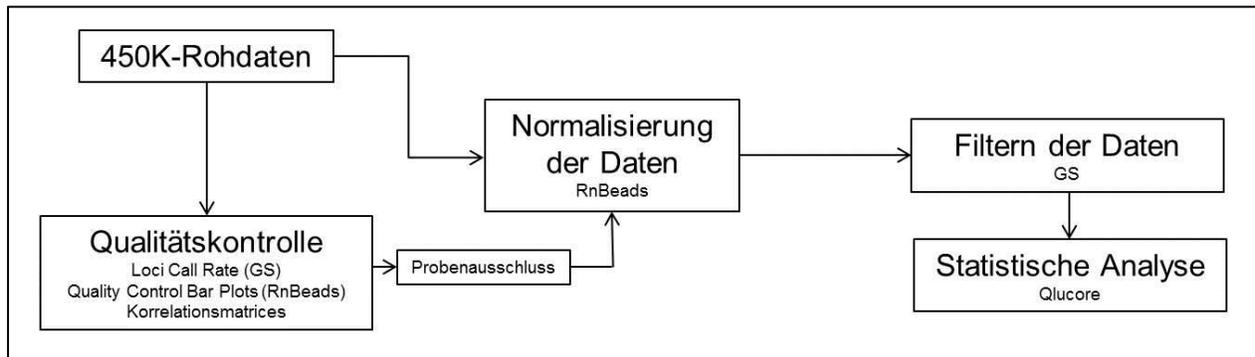


Abbildung 11 Filter-Pipeline zur Datenfilterung, Normalisierung und Qualitätskontrolle von 450K-DNA-Methylierungsdaten. GS = GenomeStudio™ Software (Version 2011.1) mit Methylation Analysis Modul (Version 1.9.0) (Illumina); RnBeads: R-Package von Assenov et al. (2014) [256] zur Analyse von 450K-Array-Daten; Qlucore: Qlucore Omics Explorer Software (Version 3.0) zur statistischen Analyse

Die 450K-Array-Daten wurden zunächst mittels GenomeStudio™ Software Version 2011.1 (Illumina; nachfolgend abgekürzt: GS-Software) mit dem GenomeStudio™ Methylation Module (Version 1.9.0, Illumina) hinsichtlich der **Hybridisierungsqualität** analysiert. Die GS-Software berechnet für jeden Locus einen *Detection Pval* (Detektions-p-Wert), welcher durch den Vergleich der Fluoreszenzsignalintensität jedes zu analysierenden Locus mit der Fluoreszenzsignalintensität Array-interner Negativkontrollsonden (zufällig permutierte, konvertierte Sequenzen, die keine CpGs enthalten und keine Hybridisierung an Proben-DNA zeigen sollten) ermittelt wird. Der Mittelwert dieser Negativkontrollsonden wird als Hintergrundsignal definiert. In dieser Arbeit wurde zur Analyse der Hybridisierungsqualität eine *Loci Call Rate* (LCR) ermittelt, definiert als der prozentuale Anteil der Loci, welche einen *Detection Pval* < 0,01 aufwiesen. Die Formel ist im Folgenden dargestellt:

$$LCR \text{ pro Probe} = \frac{\text{Anzahl Loci mit Detection Pval} < 0,01}{\text{Gesamtanzahl Loci 450K Array (485577 Loci)}} * 100$$

Mithilfe der LCR lässt sich somit der Anteil Sonden pro analysierter Probe erfassen, die ein signifikantes Fluoreszenzsignal in Relation zum Hintergrundsignal zeigen, was eine Aussage zur Hybridisierungsqualität zulässt. Als Ausschlusskriterium für Proben galt eine LCR < 98 %.

Eine weitere **Qualitätskontrolle** hinsichtlich der Bisulfit-Konversionseffizienz erfolgte durch Anwendung von RnBeads [256], einem R-*Package* zur Analyse von DNA-Methylierungsdaten mittels R-Software Version 3.1.0 (R Development Core Team). Wie bei der GS-Software können auch hier die IDAT-Dateien des 450K-Arrays eingeladen und die AVG-Beta-Werte analysiert werden. Die Anwendung des Skripts ist in Abschnitt 8 (Tabelle 33) beschrieben. Sogenannte *Quality Control Bar Plots* als Resultat der RnBeads-Analyse wurden zur Qualitätskontrolle der Bisulfit-Konversion analysiert. Die *Quality Control Bar Plots*, Balkendiagramme, ergaben Informationen über die Fluoreszenzsignalintensitäten Array-interner Bisulfit-Konversionskontrollsonden (BSK-Sonden) pro analysierter DNA-Probe. Der 450K-Array umfasst hierbei sechs BSK-Sonden des Infinium I-Assays und vier BSK-Sonden des Infinium II-Assays. Je nach BSK-Sondendesign werden im grünen oder roten Wellenlängenbereich hohe Fluoreszenzsignalintensitäten (> 2000 arbiträre Fluoreszenzeinheiten (FE)) oder niedrige Signalintensitäten (> 200 FE (Hintergrundsignal)) bei effizienter Bisulfit-Konversion erwartet. Zeigte eine analysierte DNA-Probe bspw. ein Fluoreszenzsignal > 2000 FE bei einer „unkonvertiert“-BSK-Sonde (Hintergrundsignal erwartet), so lag eine unvollständige Bisulfit-Konversion der Probe vor, sodass diese aus der weiteren Analyse ausgeschlossen wurde (siehe 3.2.2, Abbildung 25).

Als zusätzliche Qualitätskontrolle wurden zur Überprüfung von Datenhomogenität und –Qualität in dieser Arbeit zu jedem Teilprojekt sogenannte **Korrelationsmatrices** durch Verwendung der R-Software erstellt. Eine detaillierte Beschreibung zu der Erstellung der Korrelationsmatrices wird in Abschnitt 2.2.4, sowie das R-Skript in Abschnitt 8 (Tabelle 36) aufgeführt. Die Korrelationsanalyse nach Pearson-Methode der AVG-Beta-Werte pro Teilprojekt diente der Generierung von Korrelationskoeffizienten (r) für jede Probe. Als Maß für den Grad des linearen Zusammenhangs zwischen den analysierten Proben gibt der r -Wert eine Aussage über den statistischen Zusammenhang der gemessenen AVG-Beta-Werte der Proben. In dieser Arbeit wurde der r -Wert als Qualitätsparameter analysiert, um mögliche Probenverwechslungen und statistische Ausreißer identifizieren zu können. Die Ausschlusskriterien von Proben nach Analyse der r -Werte der Korrelationsmatrix werden in Abschnitt 2.2.4 beschrieben.

Im Anschluss an die Analyse der Korrelationsmatrices erfolgte eine **Normalisierung** der 450K-Array-Daten durch Verwendung des RnBeads R-*Package* und R-Befehlen zur Normalisierung. Das hierzu angewandte R-Skript ist in Abschnitt 8 anhand eines Beispiels in Tabelle 33 erläutert. Die Normalisierung der Daten dient der Vermeidung des Effekts technisch-basierter Variationen. Beispielsweise können Fluoreszenzsignal-Variationen durch Batch-Effekte auftreten. Diese können z.B. beim Vergleich von Proben, die auf unterschiedlichen Arrays und

zu unterschiedlichen Zeitpunkten prozessiert wurden, detektiert werden und würden das Ergebnis der nachfolgenden statistischen Analyse der Daten verfälschen (Harper et al., 2013). Mithilfe des R-Skripts (siehe Abschnitt 8 Tabelle 33) wurden die Fluoreszenzsignalintensitäten mittels *Subset-quantile Within Array Normalization* (SWAN) -Methode [287] normalisiert. Hierbei wird im Prinzip die Verteilung der Sonden-Fluoreszenzsignalintensitäten für jeden Array in einem Set von mehreren Arrays (hier pro Teilprojekt) angeglichen. Als Resultat der Normalisierung konnte eine Tabelle mit entsprechend normalisierten AVG-Beta-Werten der 485577 Loci pro Probe ein jedes Teilprojekts exportiert werden (hier nachfolgend bezeichnet als Normal-Tabelle).

Filtern der Daten: Parallel wurden die 450K-Array-Rohdaten mit Hilfe der GS-Software nach den unten aufgeführten Kriterien gefiltert. Zudem erfolgte eine Filterung entsprechend der Angaben von Chen et al. (2013) [288] zu kreuzreaktiven Sonden, sowie polymorphen CpGs (CpGs, die mit SNPs überlappen und einem genetischen Polymorphismus unterliegen) gefiltert. Nach der Studie von Chen et al. (2013) wurden zum einen kreuzreaktive autosomale Sonden detektiert, die eine Kohybridisierung mit DNA von Geschlechtschromosomen aufwiesen, und zum anderen konnten Sonden nachgewiesen werden, deren Ziel-CpG-Sequenz mit SNPs (1000G rel. 20110521) überlappten. Basierend auf Listen nach Chen et al. (2013) zu polymorphen CpGs, SNPs in der reaktiven Sonden-Sequenz und kreuzreaktiven Sonden wurde nach folgenden Kriterien gefiltert (im Folgenden als GS-Filterung bezeichnet):

- (1) *Detection Pval* < 0,01: nur Loci mit einem *Detection Pval* < 0,01 wurden analysiert, Loci, die dies nicht erfüllten, wurden ausgeschlossen.
- (2) Loci != ChrX/ChrY: X- bzw. Y-chromosomale Loci wurden ausgeschlossen, um geschlechtsspezifische Effekte zu vermeiden
- (3) SNPs > 0,05 AF: Ausschluss von Loci, die mit SNPs mit einer Allelfrequenz > 0,05 überlappen
- (4) SNPs > 0,05 AF in reaktiver SONDENSEQUENZ: Ausschluss von Sonden, die eine Überlappung zu SNPs mit einer Allelfrequenz > 0,05 innerhalb der Distanz von – 3 bp und + 3 bp um die Ziel-CpG-Position (Cytosin in Infinium II-Assay, Base in – 1 Position vor Ziel-CpG in Infinium I-Assay)
- (5) Ausschluss kreuzreaktiver Sonden

Anschließend wurde die gefilterte Liste der AVG-Beta-Werte der Proben mit zusätzlichen CpG-Loci-Informationen (nach GRCh37) zu *TargetID* (CpG-Loci- bzw. ch-Loci-Nummer), *CHR*, *Mapinfo* (chromosomale Position des CpG-Locus), *UCSC_RefGene_Name* (assoziiertes Gen zu dem CpG-Locus), *Relation_to_UCSC_CpG_Island* (Position des CpG-Locus relativ zu

einer CGI), *DMR* (differenziell methylierte Region), *Enhancer* und *DHS* (*DNase I hypersensitive site*) aus der GS-Software exportiert. Diese gefilterte Liste (GS-Filtertabelle) und die durch Normalisierung erstellte Normal-Tabelle (s.o.) wurden zusammengeführt, sodass eine Tabelle der gefilterten Loci mit normalisierten AVG-Beta-Werten sowie CpG-Loci-Informationen (Normal-GS-Tabelle) resultierte.

Diese Normal-GS-Tabelle wurde für die nachfolgenden finalen statistischen Analysen mittels Qlucore Omics Explorer Version 3.0 (Qlucore-Software im Folgenden abgekürzt), sowie als Hintergrundliste für funktionelle Anreicherungsanalysen (Details siehe 2.2.5.3), oder die Berechnung des *Odds Ratios* (OR) (siehe 2.2.4) verwendet. Die Methoden zu *Principal Component Analysis* (PCA), unsupervidiertes Analyse, Multivarianzanalyse (ANOVA) und T-Test für die jeweiligen DNA-Methylierungsanalysen sind unter Abschnitt 2.2.5.2 beschrieben.

2.2.2.3 PCR-basierte DNA-Methylierungsanalyse mittels Bisulfit-

Pyrosequenzierung

Die Bisulfit-Pyrosequenzierung (BSPS) ermöglicht die Quantifizierung der DNA-Methylierung von ausgewählten Cytosinen. Das Prinzip der BSPS-Methode wird im folgenden Absatz kurz zusammengefasst und anschließend die Durchführung (s.u.) genauer erläutert.

Im Vorfeld der BSPS findet zunächst eine Basen-Konversion (C zu U) der zu analysierenden DNA-Probe im Rahmen der Natriumbisulfit-Behandlung (Natriumbisulfit-Behandlung, siehe 2.2.2.1) statt. Die konvertierte DNA-Probe besteht aus zwei nicht mehr komplementären Strängen, indem konvertierte Cytosine nun als Uracil einem Guanin auf dem Komplementärstrang gegenüberstehen. Durch Verwendung von BSPS-Primern, die spezifisch an konvertierte DNA hybridisieren, erfolgt bei der BSPS-Methode im ersten Schritt eine selektive PCR-Amplifikation des konvertierten DNA-Strangs, sodass ein C/T-Polymorphismus generiert wird. Einer der beiden Primer des Primerpaares ist hierbei Biotin-markiert. Anschließend werden diese Biotin-markierten PCR-Amplifikate selektiv an Streptavidin-Sepharose-*Beads* gebunden. Es erfolgt ein Denaturierungs- und Waschvorgang dieser *Bead*-gebundenen PCR-Amplifikate. Es werden hierbei durch die Denaturierung einzelsträngige PCR-Produkte generiert, sodass nach dem Waschvorgang lediglich der *Bead*-gebundene Biotin-markierte Strang erhalten bleibt. Danach erfolgt die Hybridisierung eines weiteren Primers, des Sequenzierprimers, welcher spezifisch an einem Ende des Biotin-markierten Strangs bindet. Anschließend erfolgt die Pyrosequenzierung dieses Produkts mithilfe des PyroMark Q96 ID-Sequenzierer (Qiagen).

Durch ein zuvor definiertes Assay-Design, entsprechend der zu analysierenden DNA-Sequenz, ermittelt die PyroQ-CpG™ Software (Version 1.0.11, Qiagen), in welcher Reihenfolge die Desoxynukleotidtriphosphate (dNTPs¹⁹) – bestehend aus Desoxycytidintriphosphat (dCTP), Desoxyguanosintriphosphat (dGTP), Desoxythymidintriphosphat (dTTP) und Desoxyadenosin-Alpha-Thiotriphosphat (dATPaS²⁰) – zur Sequenzierreaktion hinzugegeben werden (*Dispension-Order*). Hierbei handelt es sich um eine *Sequencing-by-Synthesis*-Methode [289], indem die dNTPs jeweils einzeln, sowie ein Mix je aus Enzymen und Substraten, zur Reaktion hinzugefügt werden. Das Prinzip der BSPS basiert auf der Detektion von Lichtsignalen, die nach Einbau von dNTPs durch eine enzymatische Kaskade mit Freisetzung von Pyrophosphat entstehen [290]. Eine DNA-Polymerase baut zunächst das benötigte dNTP komplementär zum Ausgangsstrang ein. Das hierbei freigesetzte Pyrophosphat dient einer ATP-Sulfurylase als Co-Faktor zur Umsetzung von Adenosinphosphosulfat zu Adenosintriphosphat (ATP). Das ATP wird anschließend als Co-Faktor zur Umsetzung von Luziferin zu Oxyluziferin durch eine Luziferase genutzt. Während dieser enzymatischen Reaktion wird Energie in Form eines Lichtsignals freigesetzt, welches mithilfe des Detektors des Pyrosequenziergerätes erfasst wird. Das Lichtsignal wird als *Peak* im sogenannten Pyrogramm (Beispiel siehe Abbildung 12) dargestellt. Vor dem Hinzufügen des nächsten dNTPs zur Sequenzierreaktion werden durch eine Apyrase nicht eingebaute dNTPs abgebaut. Der Einbau zweier gleicher Nukleotide hintereinander ergibt im Pyrogramm eine doppelte *Peak*-Höhe. Bei den potenziell methylierten Loci erfolgt ein Einbau von C bzw. G (beim methylierten Locus), bzw. von T bzw. A (beim nicht-methylierten Locus). Im Pyrogramm (Abbildung 12) wird dann, je nach eingebautem C/T (bzw. G/A)-Verhältnis, ein *Peak* in entsprechender *Peak*-Höhe dargestellt. Die PyroQ-CpG™ Software (Qiagen) ermittelt dann aus der Fläche unterhalb des *Peaks* einen prozentualen DNA-Methylierungswert des entsprechenden CpG-Locus. Die Software berechnet zudem eine theoretische Lichtsignalintensität der zu analysierenden DNA-Sequenz, welche als grau hinterlegte Balken im Pyrogramm zu erkennen sind (Abbildung 12). Ein Abgleich dieser theoretisch berechneten Signalintensitäten mit den gemessenen *Peaks* dient der **Qualitätskontrolle**, indem diese möglichst übereinstimmen sollten. Liegen hier Diskrepanzen vor, so hat beispielsweise der Sequenzierprimer nicht spezifisch an die zu analysierende Sequenz gebunden. Grund hierfür können z.B. SNPs in der Sequenz sein, die somit die Hybridisierung des Primers beeinträchtigen. In diesem Fall ist ein neues Primer-Design des Sequenzierprimers erforderlich. Zusätzlich können in der *Dispension Order* in der Software-

¹⁹ dNTPs der Pyrosequenzierung nachfolgend abgekürzt als: dCTP = C, dTTP = T, dGTP = G dATPaS = A

²⁰ dATPaS wird anstelle von dATP verwendet, da dieses von der DNA-Polymerase eingebaut werden kann, der Sulfurylase jedoch nicht als Co-Faktor dient

Anwendung u.a. weitere, in der ursprünglichen Sequenz nicht enthaltene, C-Nukleotide hinzugefügt werden, die in der Sequenzierreaktion als Konversionskontrollen (gelbe Balken in Abbildung 12) dienen. Diese werden vor oder nach einem T in der *Dispension Order* hinzugefügt, um so eine mögliche unvollständige C-zu-T-Konversion eines potenziellen nicht-CpG-Cytosins (an der T-Position) detektieren zu können. Bei Detektion eines *Peaks* bei einer solchen Konversionskontrolle würde die Probe aus der Analyse ausgeschlossen werden.

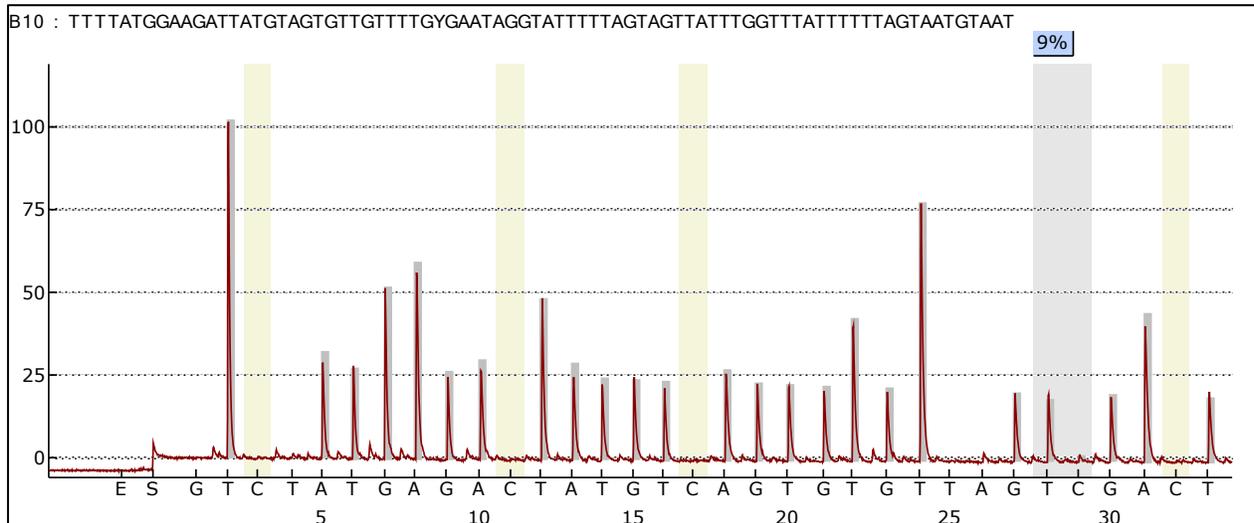


Abbildung 12 Beispiel einer Pyrogramm-Darstellung der PyroQ-CpG™ Software als Resultat der Bisulfit-Pyrosequenzierung. Dargestellt ist das Pyrogramm (Abfolge gemessener Lichtsignalintensitäten) einer zu analysierenden DNA-Sequenz mit einem CpG-Locus (durch grauen breiten Balken markiert). Über dem Pyrogramm ist die zu analysierende bisulfit-konvertierte DNA-Sequenz dargestellt. Die x-Achse zeigt die Reihenfolge der zur Sequenzierreaktion hinzugefügten dNTPs. Zuerst wird hierbei der Enzymmix (E; Enzyme für Sequenzierreaktion) und der Substratmix (S) hinzugefügt. Anschließend wird ein Test-dTNP eingespritzt, welches nicht dem nachfolgenden Nukleotid der zu analysierenden Sequenz entspricht, um die Spezifität der Reaktion zu testen (hier sollte kein *Peak* erzeugt werden). Die *Dispension Order* enthält zudem nicht-CpG-Cytosine, die nicht in der zu analysierenden konvertierten DNA-Sequenz vorkommen. Diese dienen als Konversionskontrollen (bei effizienter Bisulfit-Konversion wird kein *Peak* erwartet), welche als gelbe Balken markiert sind. Des Weiteren werden in der *Dispension Order* zusätzlich dNTPs eingefügt, die nicht Teil der zu analysierenden Sequenz sind, sodass hier keine *Peaks* gemessen werden. Grau hinterlegte Balken markieren die von der Software berechneten theoretischen Lichtsignalintensitäten der zu analysierenden Sequenz. Ein Übereinstimmen dieser theoretischen Signalintensitäten mit den gemessenen Lichtsignalintensitäten (rote *Peaks*) lässt auf einen qualitativ hochwertigen Assay schließen. Die y-Achse stellt die Lichtsignalintensität (arbiträre Einheit) dar. Der Einbau zweier gleicher Nukleotide hintereinander ergibt eine doppelte *Peak*-Höhe. Der graue breite Balken markiert die zu analysierende CpG-Position (in der zu analysierenden Sequenz als Y dargestellt) mit einer Prozent-Angabe der gemessenen DNA-Methylierung.

Die wesentlichen Vorgänge der Methode der BSPS umfassen (1) Assay-Design, (2) Test-PCRs zur Etablierung des Assays, (3) PCR der zu analysierenden Proben (4) Aufreinigung und Sequenzierprimer-Hybridisierung und (5) Pyrosequenzierung. Diese Vorgänge werden im Folgenden kurz beschrieben. Hierbei wurde das PyroMark® PCR Kit (Qiagen) und das PyroMark® Gold 96 Reagents Kit (Qiagen) verwendet.

Das **Assay-Design** mittels PyroMark® Assay Design Software (Version 2.0, Qiagen) umfasste das Primer-Design von den flankierenden PCR-Primern, sowie des Sequenzierprimers. Als

Ausgangsbasis dienten differenziell methylierte CpG-Loci, die in 450K-DNA-Methylierungsanalysen detektiert wurden und nun mittels Pyrosequenzierung verifiziert werden sollten. In die Software wurde zunächst eine ca. 400 bp lange gDNA-Sequenz (entsprechend der Sequenz aus der UCSC Genome Bioinformatics Datenbank nach GRCh37/hg19) von -200 bp vor und +200 bp nach dem zu analysierenden CpG eingeladen und durch die Software-Funktion *in silico* konvertiert. Anhand dieser konvertierten Sequenz konnte das Primer-Design entsprechend der in Tabelle 15 genannten Kriterien durchgeführt werden, wobei die Software die bestmöglichen Primerkombinationen mit Angaben zur Primerqualität darstellte. Die Primer waren hierbei spezifisch für die konvertierte DNA-Sequenz und sollten im optimierten Assay keine Hybridisierung an unkonvertierte DNA zeigen.

Tabelle 15 Kriterien für das Primer-Design mittels PyroMark® Assay Design Software (Version 2.0, Qiagen)

Kriterium	PCR-Primer	Sequenzier-Primer
Primerlänge [bp]	18 – 30 bp	15 – 25 bp
Amplifikat-Länge/ <i>TargetRegion</i> [bp]	50 - 250 bp	-
T _m min/T _m max [°C]	50 – 72 °C	29 – 59 °C
T _m -Differenz max. [°C]	10 °C	-
Max. Differenz GC-Gehalt [%]	50 %	-

T_m = Schmelztemperatur; Schmelztemperatur-Algorithmus: *Nearest Neighbour*, max = maximal; min = minimal

Für die **Primeraustestung** wurde zunächst eine Test-PCR mit fünf Kontroll-Proben durchgeführt. Dazu zählten:

- (1) konvertierte hoch-methylierte (> 85 %) Kontroll-DNA (Human high methylated genomic DNA, EpigenDx (Hopkinton, USA)), die der Definition einer maximalen Methylierungsgrenze der Pyrosequenzierung diene (MK abgekürzt)
- (2) konvertierte DNA aus einem DNA-Pool von DNA-Normalkontrollen (jeweils zehn weibliche und männliche gesunde Normalkontrollen) als Vergleichswert des „normalen“ Methylierungsstatus diene als „Normalkontrolle“
- (3) konvertierte, unmethylierte (< 5 %), Gesamt-Genom-amplifizierte DNA (WGA-DNA), welche keine DNA-Methylierung aufweist und die minimale Methylierungsgrenze definierte (WGA abgekürzt)
- (4) nicht-konvertierte DNA-Normalkontrolle, als Negativkontrolle für die PCR, um die Amplifikation nicht-konvertierter DNA ausschließen zu können
- (5) ddH₂O als Kontaminationskontrolle für die PCR

Die WGA wurde zuvor unter Verwendung von (unkonvertierter) Pool-DNA mittels Illustra GenomiPhi™ V2 DNA Amplification Kit (GE Healthcare) nach Herstellerangaben (GE Healthcare) erstellt, anschließend mittels Wizard® DNA Clean-Up System (Promega) nach

Herstellerangaben (Promega) aufgereinigt, mithilfe einer Agarose-Gelelektrophorese (siehe 2.2.1.4) überprüft und quantifiziert (siehe 2.2.1.2).

Folgender Reaktionsansatz (25 μ l) wurde standardmäßig für die PCR-Amplifikation für die Pyrosequenzierung angesetzt:

PyroMark® PCR Kit	
25 μ l-Reaktionsansatz	
12,5 μ l	PyroMark PCR Master Mix
2,5 μ l	CoralLoad
0,5 μ l	Vorwärts-Primer [20 pmol/ μ l]
0,5 μ l	Rückwärts-Primer [20 pmol/ μ l]
1 μ l	konvertierte gDNA / Kontroll-Probe
8 μ l	ddH ₂ O

Die PCR-Reaktion erfolgte nach folgendem Programm:

Denaturierung	95 °C	15 min	} 45 Zyklen
	94 °C	30 s	
Primerhybridisierung	55 - 65 °C	30 s	
Elongation	72 °C	30 s	
	72 °C	10 min	
Kühlen	4 °C	unendlich	

Die Austestung der optimalen Primerhybridisierung pro Assay wurde anhand der Hybridisierungs-Temperaturen 55 °C, 60 °C und 65 °C durchgeführt. Im Anschluss an die Test-PCR-Optimierung (nach Angaben aus Abschnitt 2.2.1.3) erfolgte eine Überprüfung der PCR-Produktlängen (5 μ l der PCR-Reaktion, CoralLoad diente direkt als Ladepuffer) mittels Agarose-Gelelektrophorese (siehe 2.2.1.4). Hierbei sollte eine distinkte intensive Bande (bemisst Effektivität der Amplifikationsreaktion) mit korrekter Amplifikat-Größe bei MK, WGA und konvertierter Pool-DNA entsprechend der Spezifität der Primer für die konvertierte DNA zu erkennen sein. Die nicht-konvertierte DNA-Normalkontrolle sowie die Kontaminationskontrolle sollten kein Produkt aufweisen. Zeigte die nicht-konvertierte DNA-Normalkontrolle dennoch ein Bandensignal, so lag eine unspezifische Hybridisierung der Primer an die unkonvertierte DNA vor und eine Optimierung des Assays oder ein neues Primer-Design war von Nöten. Sobald ein Assay optimiert vorlag, wurde dieser an Zelllinien-DNA (siehe 2.1.1.2) mit allen Kontroll-Proben

wiederholt und anschließend mit diesen PCR-Produkten die Pyrosequenzierung mit dem Sequenzier-Primer getestet. Nach Optimierung der Sequenzierreaktion (Überprüfung der Spezifität des Sequenzierprimers durch Abgleich theoretischer und gemessener Lichtsignalintensitäts-*Peaks* (s.o.)), erfolgte die PCR-Amplifikation der zu analysierenden Herzgewebeproben und den fünf Kontroll-Proben. Anschließend erfolgte wieder die Überprüfung der PCR-Produkte mittels Agarose-Gelelektrophorese (siehe 2.2.1.4).

Der PCR-Reaktionsansatz konnte anschließend unter Verwendung des PyroMark® Gold 96 Reagents Kits (Qiagen) mithilfe der PyroMark® Q96 Vacuum Workstation (Qiagen) aufgereinigt werden. Dies erfolgte nach Herstellerangaben (PyroMark Q96 MD User Manual (Qiagen)) und wird im Folgenden kurz zusammengefasst.

Durchführung: Zur Immobilisierung der biotinylierten PCR-Amplifikate (s.o.) wurde der 20 µl-PCR-Reaktionsansatz mit 20 µl ddH₂O verdünnt, in eine 96-Well Platte überführt und mit 43 µl eines Mastermixes aus 40 µl PyroMark Binding Buffer (Qiagen) und 5 µl Streptavidin Sepharose® High Performance *Beads* (GE Healthcare) versehen. Diese PCR-Platte wurde durch Abdeckfolie versiegelt und mittels Plattenschüttler (IKA® MS3 basic, IKA) für 10 min gemischt. In einer weiteren 96-Well Platte (Sequenzierplatte) wurden 40 µl PyroMark Annealing Buffer (Qiagen) und 0,5 µl Sequenzier-Primer [20 pmol/µl] gegeben und entsprechend der Herstellerangaben (PyroMark Q96 MD User Manual (Qiagen)) auf der PyroMark® Q96 Vacuum Workstation (Qiagen) platziert. Mit Hilfe des PyroMark® Q96 Vacuum Prep Tools (Qiagen) wurde die Flüssigkeit aus der Sequenzierplatte aufgenommen. Die biotinylierten, an Streptavidin-gekoppelte Sepharose-*Beads*-gebundenen PCR-Produkte werden hierbei an den Filtern des PyroMark® Q96 Vacuum Prep Tools (Qiagen) zurückgehalten. Die anschließenden Schritte zur Denaturierung und Aufreinigung an der PyroMark® Q96 Vacuum Workstation (Qiagen) erfolgten nach Herstellerangaben (PyroMark Q96 MD User Manual (Qiagen)). Hierbei erfolgt zunächst ein Waschvorgang in 70 %igem Ethanol (Details siehe 2.1.2.3) der an dem PyroMark® Q96 Vacuum Prep Tools (Qiagen) immobilisierten PCR-Produkte. Anschließend eine Denaturierung mittels 0,2 M Natriumhydroxidlösung (Details siehe 2.1.2.3), sodass die PCR-Produkte nun einzelsträngig vorlagen (Biotin-markierter Strang). Die einzelsträngigen PCR-Produkte wurden dann mittels Waschwasser (Details siehe 2.1.2.3) erneut gewaschen. Die so gereinigten, Biotin-markierten PCR-Einzelstränge wurden anschließend in die Sequenzierplatte überführt. Es erfolgte eine erneute Denaturierung bei 85 °C mittels Heizplatte (Hotplate, VWR International GmbH) und bei einer anschließenden fünf-minütigen Abkühlung auf Raumtemperatur erfolgte das Hybridisieren des Sequenzier-Primers. Die Sequenzierplatte wurde dann in dem PyroMark Q96 ID-Sequenzierer (Qiagen) positioniert, die jeweiligen Assays

(s.o.) in der PyroQ-CpG™ Software angewählt und die Pyrosequenzierung nach Herstellerangaben (PyroMark Q96 MD User Manual (Qiagen)) durchgeführt. Zur Qualitätskontrolle konnten die Pyrogramme hinsichtlich Histogramm-Peaks und Konversionskontrollen (Abbildung 12 und Beschreibung s.o.) analysiert werden.

Die **Auswertung** der erstellten DNA-Methylierungswerte erfolgte mittels Excel-Tabellenkalkulation, bzw. R-Software. Die mittels BSPS ermittelten DNA-Methylierungswerte ausgewählter CpG-Loci dienten der Verifizierung der zuvor mittels 450K-Array detektierten DNA-Methylierung der jeweiligen Proben. Zur Analyse der Korrelation der DNA-Methylierungswerte des 450K-Arrays und der BSPS wurde eine lineare Regressionsanalyse durchgeführt (siehe 2.2.4). In einem Scatterplot mit interpolierter Regressionsgerade wurden hierbei die prozentualen Methylierungswerte der BSPS auf der y-Achse gegen die 450K-Methylierungswerte (als % AVG-Beta (AVG-Beta x 100)) auf der x-Achse dargestellt. Das Bestimmtheitsmaß R^2 wurde nach Pearson-Methode ermittelt (Details siehe 2.2.4). Lag eine geringe Streuung der Messwerte um die interpolierte Regressionsgerade, sowie ein $R^2 \geq 0,9$ (aufgerundet) vor, so konnte von einer validen Verifizierung der DNA-Methylierungswerte ausgegangen werden.

2.2.2.4 PCR-basierte DNA-Methylierungsanalyse mittels MassARRAY-

EpiTYPER-Technik

Die MassARRAY-EpiTYPER-Technik (MassARRAY® EpiTYPER®, Sequenom Inc.; im Folgenden EpiTYPER-Technik abgekürzt) ist ein massenspektrometrisches Verfahren zur semi-quantitativen DNA-Methylierungsanalyse [291]. Im Vergleich zur BSPS können hier jedoch deutlich größere Amplifikate von bis zu 600 bp hinsichtlich differenziell methylierter CpGs analysiert werden und es bedarf keiner Test-PCR noch PCR-Aufreinigungen (s.u.). Im Rahmen der vorliegenden Arbeit wurde diese Technik am Institut für Humangenetik der Christian-Albrechts-Universität zu Kiel in Kollaboration mit dem Institut für Klinische Molekularbiologie (IKMB) der Christian-Albrechts-Universität zu Kiel neu etabliert. Diese Technik diene der Verifizierung von 450K-Array-Ergebnissen. Das Prinzip basiert auf einer basenspezifischen Spaltung, der *homogeneous* MassCLEAVE™ (hMC)-Reaktion (Sequenom Inc.), mit anschließender Matrix-unterstützter Laser-Desorption/Ionisation Flugzeit-Massenspektrometrie (MALDI-TOF-MS). Bei der MALDI-TOF-MS werden die zuvor generierten Spaltprodukte in einer Matrix kokristallisiert, mittels Laserimpuls herausgelöst (Desorption), in die Gasphase überführt

und ionisiert. Die ionisierten Fragmente werden dann durch ein elektrisches Feld beschleunigt. Anhand der Flugzeit kann die Masse der Fragmente bestimmt werden, indem kleinere Fragmente schneller den Detektor erreichen, als große [281]. Im folgenden Absatz werden die wesentlichen Vorgänge der EpiTyper-Technik kurz zusammengefasst und anschließend die Durchführung (s.u.) genauer erläutert. Es erfolgt:

- (1) eine PCR-Amplifikation der zu analysierenden Sequenz der konvertierten Proben-DNA, unter Verwendung T7-Promotor-gekoppelter Rückwärts-Primer (Generierung von C/T-Polymorphismus)
- (2) *Shrimp Alkaline Phosphatase* (SAP)-Behandlung zum Abbau restlicher dNTPs
- (3) *in vitro* RNA-Transkription mittels T7-RNA-Polymerase, sodass *anti-sense* RNA-Einzelstränge zur konvertierten Ausgangs-DNA entstehen (die C/T-Variation der konvertierten DNA liegt nun als G/A-Variation im RNA-Strang vor)
- (4) *homogeneous* MassCLEAVE™ (hMC)-Reaktion (Sequenom Inc.): Uracil-spezifische Spaltung der RNA mittels RNase A, sodass gleichgroße Spaltprodukte von „methylierten“ (GC) und „unmethylierten“ (AC) Sequenzen entstehen
- (5) Aufreinigung und Hybridisierung auf dem SpectroCHIP®-Array (Sequenom Inc.)
- (6) Analyse mittels MALDI-TOF-MS

Das Prinzip der EpiTyper-Technik basiert darauf, dass ein methylierter und ein unmethylierter CpG-Locus sich durch eine molekulare Massendifferenz von 16 Da unterscheiden lassen (methylierter CpG-Locus hat eine um 16 Da höhere Masse). Mittels MALDI-TOF-MS werden die Massenspektren der generierten Spaltprodukte, die die jeweiligen CpG-Loci enthalten, detektiert. Hierbei kann dann anhand des Massenspektrums die relative DNA-Methylierungsquantität durch die Berechnung der Differenz von Signalintensitäten der Massensignale von methylierten und unmethylierten *Template*-DNA-Spaltprodukten bestimmt werden. Beispielsweise resultiert ein Spaltprodukt mit einem darin enthaltenen CpG-Locus in einem messbaren Signal-*Peak*-Paar mit einer Differenz von 16 Da zueinander. Die *Peak*-Höhe von methyliertem und unmethyliertem *Peak* ist bei 50 %-Methylierung gleich hoch, während bei 100 %-Methylierung jeweils nur der methylierte *Peak* ausgeprägt ist und bei 0 %-Methylierung nur der unmethylierte *Peak* vorliegt. Das Methylierungs-*Ratio* als Beta-Wert²¹ (Werte 0 – 1, 0 = unmethyliert und 1 = voll methyliert) wird dann durch die Software (EpiTYPER™ v1.2 Software, Sequenom Inc.) dargestellt.

²¹ Dieser Beta-Wert ist nicht zu verwechseln mit dem AVG-Beta-Wert, der mittels GS-Software aus den Fluoreszenzsignalintensitäten des 450K-Arrays generiert wird. Der mittels EpiTyper-Technik erstellte Beta-Wert basiert hingegen auf der Messung von Massenspektren.

Das **Primer-Design** erfolgte mittels EpiDesigner-Software (<http://www.epidesigner.com/>, Sequenom Inc.). Die Erstellung einer Import-Datei mit einer Liste der zu analysierenden CpGs, der Position im Genom und der genomischen Sequenz (GRCh37/hg19) 200 bp *upstream* und *downstream* des CpGs erfolgte mittels R-Software und ist unter Abschnitt 2.2.5.1.1 beschrieben. Folgende Kriterien für das Primer-Design wurden in der Software als Grundlage festgelegt:

Tabelle 16 Kriterien des Primer-Designs für die EpiTyper-Technik

Kriterium	Einstellung
T _m Primer [°C]	min: 56 °C, opt.: 62 °C, max.: 64 °C
Primerlänge [bp]	min: 20 bp, opt.: 25 bp, max.: 30 bp
Amplikon-Länge [bp]	min: 100 bp, opt.: 300 bp, max.: 500 bp
Strang für Primer-Design	„beide“
Analyse CpGs in C/T Reaction	nur T reaction
Mass Window [Da]	Low: 1500 Da, High: 7000 Da

Amplikon = PCR-Amplifikat

Der Rückwärts-Primer lag nach dem Primer-Design hierbei mit einem T7-Promotor-Tag am 5'-Ende vor. Der Vorwärts-Primer wurde mit einer 10 bp-Tag-Sequenz am 5'-Ende versehen, um die Primerlängen auszubalancieren (siehe 2.1.2.1). Geeignete Amplikons (hier Bezeichnung für PCR-Amplifikate), die nach Vorhersage der EpiDesigner-Software potenziell die Detektion der zu analysierenden CpGs mittels EpiTyper-Technik ermöglichen, wurden ausgewählt und als Excel-Tabelle exportiert (R-Skript siehe Abschnitt 8 Tabelle 34). Pro zu analysierendem CpG-Locus lagen so verschiedene Amplikons des Primer-Designs vor, die anschließend mittels *in silico* Assay-Vorhersage nach Thompson et al. (2009) [255] für ein optimales Assay-Design getestet werden konnten (siehe 2.2.5.1.1, R-Skript siehe Abschnitt 8 Tabelle 35).

Durchführung

Nachdem die zu analysierenden gDNA-Proben mittels EZ DNA Methylation™ Kits (Zymo Research) konvertiert wurden (Natriumbisulfid-Behandlung, siehe 2.2.2.1), erfolgte hier im letzten Schritt der Konversion die Elution in 100 µl Wasser, sodass als Ausgangsmenge ca. 10 ng konvertierte DNA für die anschließende PCR-Amplifikation vorlag. Folgende Kontrollen wurden verwendet:

- (1) konvertierte hoch-methylierte (> 85%) Kontroll-DNA (Human high methylated genomic DNA, EpiGenDx (Hopkinton, USA)), 10 ng
- (2) unmethylierte (< 5 %) Kontroll-DNA (Human low methylated genomic DNA, EpiGenDx (Hopkinton, USA)), 10 ng
- (3) Kontaminationskontrolle (ddH₂O) für die PCR-Amplifikation

Die konvertierte Proben-DNA sowie die drei Kontrollen wurden anschließend für eine **PCR-Amplifikation** in 384-Well-Platten mit dem zuvor optimierten Assay eingesetzt.

Folgender 3 µl-PCR-Reaktionsansatz wurde dazu verwendet:

EpiTYPER® T Complete Reagent Set	
3 µl-Reaktionsansatz für 384-Well Platten	
1,42 µl	ddH ₂ O
0,5 µl	10x PCR Puffer
0,02 µl	10-mer Vorwärts-Primer [100 pmol/µl]
0,02 µl	T7-Rückwärts-Primer [100 pmol/µl]
0,04 µl	FastStart Taq DNA Polymerase (Roche Applied Science) [5 U/µl]
1 µl	konvertierte gDNA / Kontroll-Probe [10ng/µl]

Die PCR-Reaktion mittels Applied Biosystems® GeneAmp® PCR System 9700 PCR-Gerät (Life Technologies, Carlsbad, USA) erfolgte nach folgendem Programm:

Denaturierung	95 °C	2 min	} 45 Zyklen
	95 °C	20 s	
Primerhybridisierung	56 °C	30 s	
Elongation	72 °C	60 s	
	72 °C	3 min	
Kühlen	4 °C	unendlich	

Die Durchführung der PCR-Amplifikation bis hin zur Chip-Hybridisierung und MALDI-TOF-MS erfolgte nach Herstellerangaben (MassARRAY EpiTYPER v1.2 User Guide (Sequenom Inc.)) am IKMB der Christian-Albrechts-Universität zu Kiel. Im Folgenden wird das Protokoll, das den Herstellerangaben entspricht (MassARRAY EpiTYPER v1.2 User Guide (Sequenom Inc.)), kurz erläutert. Reagenzien und Puffer des EpiTYPER® T Complete Reagent Set (Sequenom Inc.) wurden verwendet. Im Anschluss an die PCR-Amplifikation erfolgte die **SAP-Reaktion** zum Abbau restlicher dNTPs. Hierzu wurde zu jedem Well 2 µl Enzymmix (0,3 µl SAP-Enzym (1,7 U/µl) und 1,7 µl ddH₂O) hinzugegeben und unter folgenden Bedingungen im Thermocycler (Applied Biosystems® GeneAmp® PCR System 9700, Life Technologies) inkubiert:

37 °C	40 min
85 °C	5 min
4 °C	unendlich

Anschließend wurde 1 µl des PCR/SAP-Produkts zur Qualitätskontrolle mittels Agarose-Gelelektrophorese (siehe 2.2.1.4) überprüft. 2 µl des PCR/SAP-Produkts wurde jeweils in eine neuen 384-Well Platte überführt. Die *in vitro* RNA-Transkription und Uracil-spezifische Spaltung mittels RNase A (hMC-Reaktion) erfolgte durch Hinzufügen von folgendem Reaktionsansatz:

EpiTYPER® T Complete Reagent Set	
5 µl- <i>MassCleave</i> -Reaktionsansatz für 384-Well Platten	
3,15 µl	ddH ₂ O
0,89 µl	5x T7 Polymerase Puffer
0,24 µl	T Cleavage Mix
0,22 µl	DTT [3,14 mM]
0,44 µl	T7 RNA & DNA Polymerase [50 U/µl]
0,06 µl	RNase A [0,09 µg/µl]

Die Reaktion wurde im Thermocycler (Applied Biosystems® GeneAmp® PCR System 9700, Life Technologies) für 3 h bei 37 °C inkubiert. Anschließend wurden die generierten Spaltprodukte mit 18 µl ddH₂O pro Well verdünnt, mittels Plattenschüttler (High-Speed Microplate Shaker, Illumina) gemischt und abzentrifugiert (Megafuge 40R, Heraeus). Dann wurde zur Aufreinigung der Spaltprodukte 6 mg CLEAN Resin (EpiTYPER® T Complete Reagent Set, Sequenom Inc.) zu jedem Well hinzugegeben und abgedeckt für 20 – 30 min rotierend gemischt. Anschließend wurde für 5 min bei 3200 g zentrifugiert (Megafuge 40R, Heraeus). Die Beladung (Dispension) der SpectroCHIP®-Arrays (EpiTYPER®, Sequenom Inc.) erfolgte mittels MassARRAY™ Nanodispenser RS1000 (Sequenom Inc.) nach Herstellerangaben (MassARRAY Nanodispenser RS1000 User Guide, Sequenom) mit einer Dispensionsgeschwindigkeit von 110 mm/s und 0,2 s Dispensionszeit. Die Messung mittels MassARRAY™ Analyzer Compact (Sequenom Inc.), einer MALDI-TOF-MS-Technologie, wurde nach Herstellerangaben (MassARRAY Analyzer Compact Users Guide (Sequenom, Inc)) durchgeführt. Dazu wurden zuvor die Probenbezeichnungen und Assays in das Modul „EpiTYPER Plate Editor“ der EpiTYPER™ v1.2 Software (Sequenom Inc.) eingegeben. Im Anschluss an die MALDI-TOF-MS-Messung wurden dann mittels „EpiTYPER Analyzer“-Modul (EpiTYPER™ v1.2 Software, Sequenom Inc.) die DNA-Methylierungswerte als Beta-Werte anhand des Massenspektrums berechnet.

Qualitätskontrolle

Für eine valide Auswertbarkeit musste die Spektrumqualität pro Assay für jede Probe eine *Confidence* (Reliabilitäts-Ranking) von 4 – 5 nach „EpiTYPER Analyzer“-Modul (EpiTYPER™ v1.2 Software, Sequenom Inc.) haben (0 = *low confidence*, 5 = *high confidence*). Weitere Kriterien zur Auswertbarkeit des zu analysierenden CpGs werden in Tabelle 17 genannt. Bei mehr als einem CpG pro Fragment (Spaltprodukt) werden von der Software gewichtete Mittelwerte der DNA-Methylierung aller CpGs dieser sogenannten CpG Unit berechnet. In dieser Arbeit wurde ein Maximum von 2 CpGs pro Unit als auswertbar definiert (nach „Sequenom Product Preview Note“ [291]), um eine Verfälschung der Messergebnisse zu vermeiden. Zudem können CpGs in Fragmenten liegen, deren Masse außerhalb des messbaren Bereichs von 1500 – 7000 Da ist (*low mass/high mass*), sodass hier keine Auswertung möglich war. Wenn Fragmente derselben Masse (überlappende Fragmente) vorliegen, werden diese als SM (*silence methylated*)- und SN (*silence nonmethylated*)-Fragmente von der Software annotiert. Im Falle eines SM-Fragments liegt zusätzlich zu dem Analyt-Fragment, das eine CpG-Methylierung aufweist, eine Überlappung mit einem Nicht-CpG-Fragment vor, das dieselbe Masse hat. Bei SN-Fragmenten kann eine Überlappung des nicht methylierten Analyt-Fragments zu einem Nicht-CpG-Fragment derselben Masse detektiert werden. Hierbei wurde in dieser Arbeit ein Maximum von einem SM-/SN-Fragment als Kriterium für die Auswertbarkeit des CpGs definiert (Tabelle 17). Erfüllte das zu analysierende CpG die Kriterien zur Auswertbarkeit, konnten die Beta-Werte zu dem jeweiligen Assay als Excel-Tabelle exportiert und zum Abgleich mit den 450K-DNA-Methylierungswerten verwendet werden.

Tabelle 17 Kriterien zur Auswertbarkeit des zu analysierenden CpGs bei Analyse mit dem „EpiTYPER Analyzer“-Modul (EpiTYPER™ v1.2 Software, Sequenom Inc.)

Parameter	Kriterium
CpG Unit	≤ 2 CpGs
L_mass (<i>low mass</i>) / H_mass (<i>high mass</i>)	keine Auswertung möglich
SM (<i>silence methylated</i>)	< 2
SN (<i>silence nonmethylated</i>)	< 2

CpG Unit: CpGs pro Fragment (bei mehr als 1 CpG pro CpG Unit werden gewichtete Mittelwerte der DNA-Methylierung aller CpGs der CpG Unit dargestellt); SM: mit dem zu analysierenden Fragment, das methyliert vorliegt (analytisches Fragment), liegt eine Überlappung eines anderen Nicht-CpG-Fragments vor, welches dasselbe Molekulargewicht hat und somit gleichsam detektiert wird; SN: mit dem zu analysierenden Fragment, das nicht methyliert vorliegt (analytisches Fragment), liegt eine Überlappung eines anderen Nicht-CpG-Fragments vor, welches dasselbe Molekulargewicht hat und somit gleichsam detektiert wird.

Die **Auswertung** der erstellten DNA-Methylierungswerte (Beta-Werte von 0 bis 1) erfolgte mittels Excel-Tabellenkalkulation bzw. R-Software. Die mittels EpiTyper-Technik ermittelten DNA-Methylierungswerte ausgewählter CpG-Loci dienen der Verifizierung der zuvor mittels 450K-Array detektierten DNA-Methylierung der jeweiligen Proben. Zur Analyse der Korrelation

der DNA-Methylierungswerte des 450K-Arrays und der Epityper-Technik erfolgte eine lineare Regressionsanalyse (siehe 2.2.4). Hierbei wurden in einem Scatterplot mit interpolierter Regressionsgerade die Methylierungswerte der Epityper-Analyse auf der y-Achse gegen die 450K-Methylierungswerte auf der x-Achse dargestellt. Das Bestimmtheitsmaß R^2 wurde nach Pearson-Methode ermittelt (Details siehe 2.2.4). Eine valide Verifizierung der DNA-Methylierungswerte war mit $R^2 \geq 0,9$ (aufgerundet) definiert.

2.2.3 Molekulargenetische Analysen von RNA

Im folgenden Abschnitt werden die Methoden zu molekulargenetischen Analysen von RNA beschrieben. In Abbildung 13 ist dies schematisch als Übersicht dargestellt.

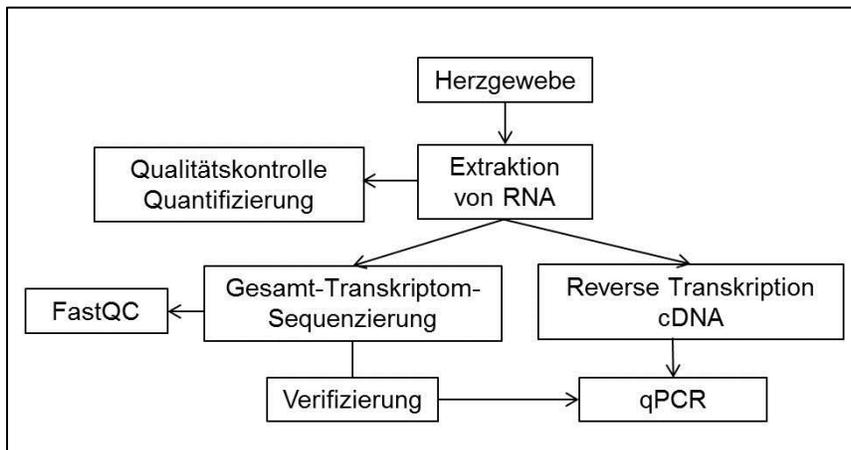


Abbildung 13 Schematische Darstellung der Methoden zu molekulargenetischen Analysen von RNA. FastQC: Programm zur Qualitätskontrolle der Rohdaten der Gesamt-Transkriptom-Sequenzierung (FastQC Version 0.11.2, Babraham Bioinformatics); cDNA: *complementary DNA*; qPCR: *quantitative Real-Time PCR*

2.2.3.1 Extraktion von RNA

Die Extraktion von RNA aus IAS-Gewebeproben wurde im Rahmen der Gesamt-Transkriptom-Sequenzierung durch AROS Applied Biotechnology A/S (Aarhus, Dänemark) durchgeführt. Hierbei wurde die Extraktion von RNA nach Standardarbeitsanweisungen (siehe <http://arosab.com/services/extraction/mrna-2/>, AROS Applied Biotechnology A/S) durchgeführt, wobei TRIzol® Reagent (Life Technologies) und Säulenaufreinigung mittels RNeasy Kit (Qiagen) angewendet wurde. Nach der Durchführung der Gesamt-Transkriptom-Sequenzierung standen restliche extrahierte RNA-Proben für nachfolgende Analysen zur Verfügung.

2.2.3.2 Quantifizierung von RNA

Die Quantifizierung der RNA der IAS-Gewebeproben erfolgte durch AROS Applied Biotechnology A/S (Arhus, Dänemark) mittels Nanodrop (ND-1000 Spektrophotometer) durch Messung des A260/A280 nm Wellenlängenverhältnisses. Ein Wert von 1,9 – 2,1 wurde als qualitativ hochwertige RNA eingestuft (siehe <http://arosab.com/services/extraction/mrna-2/>, AROS Applied Biotechnology A/S). Die Integrität wurde mittels Agilent BioAnalyzer ermittelt.

2.2.3.3 Gesamt-Transkriptom-Sequenzierung

Die Gesamt-Transkriptom-Sequenzierung (RNA-Sequenzierung im Folgenden abgekürzt) ermöglicht die Sequenzierung des gesamten Transkriptoms, d.h. die Gesamtheit aller Transkripte in einer Zelle zu einem bestimmten Entwicklungsstadium und physiologischen Bedingungen [292]. Die RNA-Sequenzierung basiert auf der Erstellung einer cDNA-*Library* (cDNA: *complementary* DNA) und Sequenzierung der cDNA-Fragmente durch Hochdurchsatz-Sequenzierungstechnologien. Im Anschluss an ein Sequenzalignment zur humanen Genom-Referenz (GRCh37/hg19) können Genexpressionslevel bestimmt und ausgewertet werden.

Die RNA-Sequenzierung erfolgte durch die Firma AROS Applied Biotechnology A/S (Arhus, Dänemark) an aus IAS-Gewebeproben extrahierter RNA. Hierbei wurde das TruSeq RNA Library Preparation Kit (Illumina) zur Erstellung der cDNA-*Library* verwendet und im Folgenden kurz erläutert. Anhand des Poly(A)-3'-Endes der mRNA-Moleküle werden diese mittels Oligo-dT-Magnet-*Beads* selektiert und aufgereinigt, anschließend fragmentiert und mittels randomisierter Hexamer-Primer und reverser Transkriptase in cDNA transkribiert. Eine Zweitstrangsynthese erfolgt dann mittels DNA-Polymerase I und RNase H. Wie bei der WES-Technik (siehe 2.2.1.7) erfolgt auch hier ein *End Repair*, *dA-Tailing* und eine Adapterligation. Die Produkte werden erneut angereichert und mittels PCR amplifiziert, sodass die erstellte cDNA-*Library* anschließend mittels Illumina HiSeq2000 System sequenziert werden konnte. In der vorliegenden Arbeit wurden die erhaltenen Rohdaten anschließend zur Qualitätskontrolle mittels FastQC Software (Version 0.11.2, Babraham Bioinformatics) analysiert (Details siehe 2.2.5.4), um z.B. Kontaminationen durch Sequenz-Adapter identifizieren zu können.

Die weitere Qualitätsfilterung und bioinformatische Auswertung erfolgte durch PhD Gabriel Musso (Brigham and Women's Hospital, Harvard Medical School, Boston) und umfasste:

- (1) *Cassava Chastity* Filterung (Illumina), um eine äquivalente *Read*-Qualität zu erhalten
- (2) Entfernung von Adaptersequenzen mittels CutAdapt [293]
- (3) Sequenz-Alignment mittels TopHat2 [294] nach Genom-Referenz GRCh37/hg19
- (4) Expressionsanalyse mittels EdgeR [254] und *SNP-Calling*

Als Ergebnis der RNA-Sequenzierung resultierten sogenannte *Counts*-Datensätze, die die Anzahl der *Reads* (sequenzierte cDNA-Fragmente) pro Gen bzw. Exon darstellen. Die zur Verfügung gestellten Daten zur differentiellen Expression von HLHS-Proben und Kontrollproben (Rechtsherzfehlbildungen), sowie der Ergebnisse des *SNP-Callings* konnten in dieser Arbeit für integrative Analysen genutzt werden.

2.2.3.4 Reverse Transkription

Die reverse Transkription, d.h. die Synthese von cDNA aus RNA, wurde mittels QuantiTect® Reverse Transcription Kit (Qiagen) durchgeführt. Pro Probe wurde je 1 µg RNA eingesetzt und nach Herstellerangaben (QuantiTect® Reverse Transcription Handbook, Qiagen) transkribiert. Dazu wurde zunächst restliche gDNA entfernt, indem 2 µl gDNA Wipeout Buffer (Qiagen) zu der auf ein 12 µl-Volumen verdünnten RNA-Probe hinzugefügt und anschließend für 2 min bei 42 °C inkubiert wurde. Anschließend wurde die Probe auf Eis gegeben und für die cDNA-Synthese in einem 20 µl-Reaktionsansatz in 200 µl Reaktionsgefäßen eingesetzt:

QuantiTect® Reverse Transcription Kit	
20 µl Ansatz	
1 µl	Quantiscript Reverse Transkriptase
4 µl	5x Quantiscript RT Puffer
1 µl	RT Primer Mix
14 µl	<i>Template</i> -RNA aus der gDNA Wipeout Reaktion

Die cDNA-Synthese erfolgte bei 42 °C für 30 min mit anschließender Inaktivierung bei 95 °C für 3 min im Thermocycler (UNO-Thermoblock, Biometra). Die cDNA konnte bei 4 °C bzw. bei -20 °C für Langzeitlagerung aufbewahrt werden. Für die nachfolgende quantitative Real-Time PCR wurde die erstellte cDNA im 1:10-Verhältnis verdünnt.

2.2.3.5 Quantitative Real-Time PCR

Die quantitative Real-Time PCR (qPCR) ist eine fluoreszenz-basierte PCR-Methode, bei der eine Echtzeit-Messung der Fluoreszenzsignale während der PCR-Zyklen erfolgt und somit eine Amplifikation und simultane Quantifizierung von Zielsequenzen ermöglicht wird. In dieser Arbeit wurde die qPCR zur Verifizierung von RNA-Sequenzierungsergebnissen angewendet. Die Patienten-RNA wurde dazu zunächst in cDNA umgewandelt (reverse Transkription, siehe 2.2.3.4) und anschließend für die qPCR-Reaktion verwendet. Als Fluoreszenzfarbstoff diente hierbei SYBR® Green (QuantiTect® SYBR® Green PCR Kit, Qiagen). Dieser basen-interkalierende Cyanin-Farbstoff bindet doppelsträngige DNA (die Proben-cDNA) und der so gebildete SYBR® Green-DNA-Komplex emittiert nach Anregung durch blaues Licht ($\lambda_{\text{max}} = 494 \text{ nm}$) ein grünes Lichtsignal ($\lambda_{\text{max}} = 521 \text{ nm}$), welches durch das LightCycler® 480 System (LightCycler® 480 II, Roche) detektiert werden kann. Die erhaltenen Fluoreszenz-Signale konnten anschließend mithilfe der Software Lightcycler®480 SW (Version 1.5.1, Roche) ausgewertet werden, um die ursprünglich eingesetzte Proben-cDNA zu quantifizieren.

Primer-Assays für die qPCR wurden kommerziell von Qiagen (Hilden, Deutschland) erworben. Die pPCR-Primer sind hierbei in der Weise erstellt, dass diese ausschließlich spezifisch an kodierende DNA-Sequenzen (cDNA), d.h. Sequenzen ohne Introns, binden. Durch das Überspannen der Primersequenz bspw. über benachbarte Exons, sind diese qPCR-Primer cDNA-spezifisch. Das Einsetzen von gDNA in der qPCR-Reaktion ergibt im optimierten Assay keine Amplifikation, da aufgrund der in der gDNA enthaltenen Intronsequenzen eine inkomplette Basenkomplementarität der qPCR-Primer besteht. DNA-Normalkontrollen (gDNA) wurden somit als Negativkontrollen verwendet.

Folgender qPCR-Ansatz (20 μl) wurde mittels QuantiTect® SYBR® Green PCR Kit (Qiagen) standardmäßig angesetzt:

QuantiTect® SYBR® Green PCR Kit	
20 μl Ansatz	
10 μl	2x QuantiTect SYBR Green PCR Master Mix
2 μl	QuantiTect Primer Assay
7 μl	ddH ₂ O
1 μl	cDNA

Je nach Verdünnung der cDNA wurden 1 -2 µl der zu analysierenden cDNA jeweils als Triplikate eingesetzt und die ddH₂O-Menge angepasst. Herz-cDNA von Clontech Laboratories Inc. diente als Positivkontrolle. Als Negativkontrolle wurde eine DNA-Normalkontrolle verwendet. Zur Kontrolle hinsichtlich Kontaminationen wurde ddH₂O in der qPCR verwendet.

Die qPCR-Reaktion wurde unter folgenden Bedingungen durchgeführt:

Denaturierung	95 °C	15 min	} 45 Zyklen
	95 °C	30 s	
Primerhybridisierung	58 °C	30 s	
Elongation	72 °C	30 s	
Schmelzkurve	72-95 °C	unendlich	

Zur **Qualitätskontrolle** wurde im Anschluss an die qPCR-Reaktion der 20 µl-qPCR-Ansatz mittels Agarose-Gelelektrophorese (siehe 2.2.1.4) überprüft. Die Negativ- und Kontaminationskontrolle sollten hierbei keine Bande ergeben, und die Proben-cDNA sowie die Positivkontrolle sollten eine distinkte Bande ergeben, damit der Assay auswertbar ist. Eine Analyse der Schmelzkurve, welche zum Ende des PCR-Programms generiert wird, diente als weitere Qualitätskontrolle. Das Prinzip basiert darauf, dass jede dsDNA bei einer spezifischen Schmelztemperatur zu einzelsträngigen Molekülen denaturiert wird und somit verschiedene PCR-Produkte unterschieden werden können. Primerdimere und unspezifische PCR-Produkte konnten so detektiert werden.

Die **Quantifizierung** der ursprünglich eingesetzten cDNA erfolgte nach Bustin et al. (2009) [295] durch die sogenannte $\Delta\Delta C_q$ -Methode, indem der Quantifizierungs-Zyklus-Wert (C_q-Wert), sowie die PCR-Effizienz einberechnet wurden. Die Primereffizienz E, die die Amplifikationsrate pro Zyklus angibt und in einer idealen PCR einer Verdopplung (E = 2) entspräche, unterscheidet sich für jedes Primerpaar und Ziel-DNA. Daher wurden Verdünnungsreihen mit Herz-cDNA von 10 ng bis 500 ng für jeden Assay in Triplikaten durchgeführt und Kalibrierkurven erstellt, indem pro Primerpaar die C_q-Werte gegen den Logarithmus der cDNA-Konzentration aufgetragen wurde. Mithilfe der ermittelten Steigung konnte anschließend die Effizienz über die Formel $E = 10^{-1/\text{Steigung}}$ berechnet werden (Grenzwert $E \geq 1.8$). Die Genexpression konnte dann unter Einberechnung von Referenzgenen (sogenannte *Housekeeping*-Gene, welche unabhängig von

Zelltyp und Zellstadium konstitutiv exprimiert werden), wobei hier *HPRT1* und *G6PD* verwendet wurden, ermittelt werden. Hierzu wurde folgende Formel verwendet:

$$\text{Ratio} = \frac{E(\text{Referenzgen})^{Cq(\text{Referenzgen})}}{E(\text{Zielgen})^{Cq(\text{Zielgen})}}$$

Von den *Ratio*-Werten pro Referenzgen wurde der Mittelwert gebildet. Die zu analysierenden Proben wurden in Duplikaten eingesetzt. Die Ergebnisse der RNA-Sequenzierung konnten dann zur Verifizierung mit denen der qPCR verglichen und mittels R-Software analysiert werden.

2.2.4 Statistik

In diesem Abschnitt werden die in dieser Arbeit angewendeten statistischen Testverfahren beschrieben. Im Folgenden werden statistische Tests und deren Parameter aufgeführt, die in den verschiedenen Teilprojekten der DNA-Methylierungsanalysen bzw. der Transkriptom-Analyse angewendet wurden. In Tabelle 18 ist eine Übersicht der angewendeten Testverfahren und Parameter aufgeführt.

Tabelle 18 Statistische Testverfahren, Parameter und Anwendung

Test	Parameter	Anwendung
<i>Student's T-Test</i> (T-Test)	$q \leq 1 \times 10^{-6}$	Zwei-Gruppen-Vergleich, allgemein
ANOVA	$q \leq 1 \times 10^{-6}$	Vergleich mehrerer Gruppen
Rangregression	$q \leq 3 \times 10^{-4}$	Variable-abhängige Analyse
Welch <i>Two Sample T-Test</i> (wT-Test)	$p \leq 0,05$	Zwei-Gruppen-Vergleich kleiner Gruppengrößen ($N \leq 5$), CpG-spezifisch
Wilcoxon <i>Signed-Rank Test</i>	$p \leq 0,05$	Zwei-Gruppen-Vergleich (große Gruppengröße)

q: *False Discovery Rate* (FDR); allgemein: bezogen auf Gesamtvergleich der DNA-Methylierung aller CpG-Loci zwischen zwei Gruppen; CpG-spezifisch: bezogen auf den Vergleich der DNA-Methylierung spezifischer CpG-Loci zwischen zwei Gruppen

In dieser Arbeit wurden die resultierenden Daten der 450K-Array-Analyse (AVG-Beta-Werte) nach vorausgegangener Normalisierung des Datensatzes (siehe 2.2.5.1.1) mittels parametrischer Statistik analysiert, auch wenn die Datensätze mathematisch gesehen keine Normalverteilung, die Grundannahme parametrischer Testverfahren, aufzeigten. Die Anwendung von parametrisch-basierten Testverfahren, insbesondere bei Auswertung von AVG-Beta-Werten, ist eine in der Literatur anerkannte Methode [296], daher wurden auch in dieser Arbeit parametrisch-basierte Testverfahren zur Analyse der 450K-Daten angewendet.

Ein **Student's T-Test** (im Folgenden T-Test abgekürzt) wurde mittels Qlucore-Software (Details siehe 2.2.5.2) zur Analyse der differentiellen DNA-Methylierung zwischen zwei verschiedenen Gruppen angewendet. Die *False Discovery Rate* (FDR, q-Wert) anstelle des p-Werts wurde hier grundsätzlich verwendet, da diese für multiples Testen korrigiert ist [297]. Als minimale Signifikanzgrenze galt ein q-Wert von $\leq 0,05$. Differentiell methylierte CpG-Loci mit $q > 0,5$ wurden als nicht signifikant erachtet. Bei Zwei-Gruppen-Vergleichen der verschiedenen Teilprojekte der DNA-Methylierungsanalysen wurde grundsätzlich ein Grenzwert von $q \leq 1 \times 10^{-6}$ verwendet²². Loci die dies erfüllten, wurden als signifikant differentiell methyliert betrachtet. Eine **Varianzanalyse (ANOVA)** wurde mittels Qlucore-Software zur Analyse der differentiellen DNA-Methylierung in (mehr als zwei) Gruppen verschiedener Herzgewebetypen mit $q \leq 1 \times 10^{-6}$ angewendet. Zur Analyse der altersabhängigen DNA-Methylierung wurde eine **Rangregressionsanalyse**, hier mit $q \leq 3 \times 10^{-4}$, durchgeführt. Zur spezifischen Analyse der differentiellen DNA-Methylierung bestimmter CpG-Loci zwischen zwei Gruppen wurde der **Welch Two Sample T-Test** (wT-Test abgekürzt) mittels R-Software angewendet, um Effekte durch ungleiche Intragruppenvarianzen und insbesondere kleiner Gruppengrößen einzuberechnen. Da es sich bei den analysierten Proben um schwer zugängliches primäres Herzgewebe handelte, war die Probenanzahl pro analysierte Gruppe teilweise zu gering, um ein anderes Testverfahren anzuwenden. Als Signifikanzgrenze galt hier ein p-Wert $\leq 0,05$. Zur Analyse differentiell exprimierter Gene zwischen zwei Gruppen anhand von *Counts*-Daten der RNA-Sequenzierung wurde das Testverfahren **Fisher's-Exact Test** angewendet. Zur Analyse von *Counts*-Daten ist letzteres Testverfahren eine in der Literatur anerkannte Methode [298]. Bei Analysen mit RVIS-Score (*Residual Variation Intolerance Score*; siehe 2.1.3.2), wie in der Gesamt-Exom-Sequenzierungsanalyse in Abschnitt 3.2.1.1 erfolgt, wurde der **Wilcoxon Signed-Rank Test** angewendet.

Bei Korrelationsanalysen fand das parametrisch-basierte Testverfahren der **Pearsons's Korrelation** mithilfe der R-Software Anwendung. Die Analyse des Korrelationskoeffizienten (r)²³ diente im Rahmen der Qualitätskontrolle bei 450K-Array-Analysen der Überprüfung von Datenhomogenität und -Qualität. Hierbei wurden in einer CpG-Gesamt-Korrelation²⁴ und einer SNP-Korrelation (rs-Loci-Korrelation) der AVG-Beta-Werte Korrelationskoeffizienten nach Pearson-Methode ermittelt. Die Erstellung der Korrelationsmatrices mittels R-Software ist in Abschnitt 2.2.5.1.2 (R-Skript siehe Abschnitt 8 Tabelle 36) beschrieben. Bei Proben desselben

²² wenn nicht anders angegeben

²³ In dieser Arbeit ist bei Nennung des r-Werts in DNA-Methylierungsanalysen (siehe Abschnitt 3.2.2) grundsätzlich der Median des r-Werts gemeint.

²⁴ Begriff schließt ch-Loci mit ein

Gewebetyps (z.B. IAS-Gewebe) werden in der **CpG-Gesamt-Korrelation** homogen hohe r -Werte erwartet. Je näher dieser bei dem Wert 1 liegt, desto höher ist die Wahrscheinlichkeit eines linearen Zusammenhangs der gemessenen AVG-Beta-Werte der Proben zueinander. Zielte die Teilprojektanalyse auf einen T-Test ab, so erfolgte ein Probenausschluss, dessen Bedingungen im Folgenden erklärt werden. Proben desselben Gewebetyps mit einem Median über alle Korrelationskoeffizienten von $r < 0,97$ wurden aus der Analyse ausgeschlossen (siehe 3.2.2, Abschnitt „Qualitätskontrolle“ pro Teilprojekt). Dieser Grenzwert von $r < 0,97$ wurde definiert durch den Ausschluss entsprechend des r -Werts der unteren 10 % der Proben anhand der initialen Analyse von Teilprojekt P1 (Details siehe Abschnitt 3.2.2.1.1). Auch Proben derselben Gewebegruppe wurden im Kontext eines nachfolgenden T-Tests nach einem Grenzwert von $r < 0,97$ ausgeschlossen. Solche Gewebegruppen vereinten jeweils verschiedene Gewebetypen (z.B. Gruppe „Atrium“ und „Ventrikel“, siehe 3.2.2.3) und wurden im Rahmen des jeweiligen Teilprojekts definiert. Proben einer Gewebegruppe mit $r < 0,97$ wurden auch hier aus der Analyse ausgeschlossen, um die Intragruppen-Heterogenität für den nachfolgenden T-Test-Vergleich dieser Gruppen zu minimieren. Schloss die Analyse lediglich Proben eines einzigen Individuums ein (wie in Abschnitt 3.2.2.4), so erfolgte kein Probenausschluss nach Korrelationskoeffizient. Wurden verschiedene Gewebetypen im Hinblick auf differenzielle DNA-Methylierungsmuster unterschiedlicher anatomischer Geweberegionen im Rahmen einer Multivarianzanalyse (ANOVA, siehe 2.2.4 und 2.2.5.2) untersucht, so erfolgte kein Ausschluss von Proben nach dem Korrelationskoeffizient-Grenzwert. Eine **SNP-Korrelationsmatrix** im Rahmen der Qualitätskontrolle bei 450K-Array-Analysen wurde anhand der AVG-Beta-Werte der 65 rs-Loci (SNPs) des 450K-Arrays nach Pearson-Methode erstellt. Bei der SNP-Korrelation sollten Proben eines Individuums, d.h. biologische oder technische Replikate bzw. verschiedene Herzgewebetypen eines Individuums, als Einschlusskriterium für nachfolgende Analysen einen Wert von r (Median) $> 0,99$ zeigen (siehe 3.2.2, Abbildung 26). Eventuelle Probenverwechslungen wären über diese SNP-Korrelation entsprechend festzustellen.

Eine **lineare Regressionsanalyse** wurde durchgeführt, um die Korrelation der mittels 450K-Array generierten DNA-Methylierungswerte und der DNA-Methylierungswerte, die mittels einer der beiden PCR-basierten Techniken, BSPS oder EpiTyper-Technik, generiert wurden, zu analysieren. Auch die Korrelation der Methylierungswerte von technischen und biologischen Replikaten wurde mithilfe der linearen Regressionsanalyse untersucht. Bei Verifizierungen wurden jeweils die DNA-Methylierungswerte des 450K-Arrays auf der x -Achse gegen die DNA-Methylierungswerte je von der BSPS oder der EpiTyper-Technik auf der y -Achse in einem

Scatterplot mit interpolierter Regressionsgerade dargestellt. Nach Pearson-Methode wurde das Bestimmtheitsmaß R^2 berechnet. Der R^2 -Wert (Werte von 0 – 1) dient hierbei als Maß für den Grad des linearen Zusammenhangs der mittels zweier verschiedener Techniken erhobenen DNA-Methylierungswerte. Eine hohe Korrelation der Daten wurde in dieser Arbeit mit $R^2 \geq 0,9$ definiert. Lag zudem eine geringe Streuung der Messwerte um die interpolierte Regressionsgerade vor, so konnte von einer validen Verifizierung der DNA-Methylierungswerte ausgegangen werden.

Das **Quotientenverhältnis (Odds Ratio (OR))**, zur Analyse der Signifikanz bzw. der Stärke des Zusammenhangs von zwei Merkmalen, wurde in dieser Arbeit zur Analyse der Anreicherung regulatorischer Elemente innerhalb differenziell methylierter CpGs verschiedener 450K-Array-Analysen angewendet. Die entsprechenden p-Werte wurden mittels *Fisher's Exact Test* berechnet. Die normalisierte und gefilterte Hintergrundliste der AVG-Beta-Werte (siehe 2.2.2.1), wurde als Referenzliste verwendet.

2.2.5 Bioinformatik

Dieser Abschnitt beschreibt die in dieser Arbeit angewendeten Methoden der Bioinformatik. Dies umfasste die bioinformatische Auswertung mittels R-Software und Qlucore-Software, sowie funktionelle Anreicherungsanalysen mittels WEB-based GENE SeT AnaLysis Toolkit (WebGestalt [278], (siehe 2.1.3.2). Des Weiteren wurden Rohdaten der RNA-Sequenzierung mittels FastQC-Software (Babraham Bioinformatics) analysiert.

2.2.5.1 Bioinformatische Auswertung mittels R-Software

In dieser Arbeit wurden Analysen mittels R-Software Version 3.1.0 (R Development Core Team) für verschiedene Teilprojekte durchgeführt, wobei teilweise mit *R-Packages* gearbeitet oder angepasste R-Skripte eigens geschrieben wurden. Die Angaben zu URL, Zitierung und Referenz der jeweiligen *R-Packages* können aus der Tabelle 11 in Abschnitt 2.1.3.1 entnommen werden.

2.2.5.1.1 R-*Package*-basierte Analysen mittels R-Software

Das **RnBeads-*Package*** [256] wurde zur Qualitätskontrolle hinsichtlich der Bisulfit-Konversion und zur Normalisierung der Daten der Array-basierten DNA-Methylierungsanalyse (450K-Array) verwendet (siehe 2.2.2.2). Hierbei werden die IDAT-Dateien (Rohdaten 450K-Array) und eine Annotationstabelle, die die Array-Positionen der Proben definiert, importiert, sodass anschließend (A) die RnBeads-Analyse zur Qualitätskontrolle, bzw. (B) die Normalisierung durchgeführt werden konnte. Die R-Befehle des R-Skripts werden in Abschnitt 8 anhand eines Beispiels in Tabelle 33 kurz erläutert. Als Resultat der RnBeads-Analyse wurden zur Qualitätskontrolle der Rohdaten des 450K-Arrays sogenannte *Quality Control Bar Plots* als Resultat der RnBeads-Analyse [256] analysiert (Details siehe 2.2.2.2.1).

Für das Primer-Design der Epi typer-Technik (siehe 2.2.2.4) war zum einen die Erstellung einer Import-Datei durch Verwendung der R-basierten Software **Bioconductor (Bioclite-Tool)** [252] für das primäre Primer-Design mittels EpiDesigner-Software notwendig, und zum anderen mussten anschließend die mittels EpiDesigner-Software ermittelten möglichen Primerkombinationen (Amplikon-Sequenzen) in einer *in silico* Fragmentierungs-Analyse mittels R-*Package* MassArray [255] analysiert werden. Für das Primer-Design der Epi typer-Technik (siehe 2.2.2.4) erfolgte somit zunächst die Erstellung einer Liste der zu analysierenden CpGs. Anschließend erfolgte anhand dieser CpG-Liste mithilfe der R-basierten Software Bioconductor die Erstellung einer Import-Datei für die EpiDesigner-Software. In Abschnitt 8 (Tabelle 34) werden die jeweiligen R-Befehle aufgeführt. Das Resultat war dann eine Liste bestehend aus den zu analysierenden CpGs, der Position im Genom und der genomischen Sequenz (GRCh37/hg19) je 200 bp *upstream* und *downstream* des CpGs.

Die mittels EpiDesigner-Software ermittelten möglichen Primerkombinationen (Amplikon-Sequenzen) wurden anschließend mittels **R-*Package* MassArray** in einer *in silico* Assay-Vorhersage analysiert, um geeignete Assays für die Epi typer-Technik auszuwählen. Die jeweiligen R-Befehle dieser *in silico* Fragmentierungs-Analyse werden in Abschnitt 8 anhand eines Beispiels in Tabelle 35 erläutert. Mithilfe dieser *in silico* Assay-Vorhersage konnte die theoretische molekulare Masse der Spaltprodukte berechnet werden. Dies diente der Vorhersage der zu erwartenden MALDI-TOF-*Peaks*, die für eine Analysierbarkeit im messbaren Bereich von 1500 – 7000 Da liegen mussten. Auch Vorhersagen für Fragmente derselben Masse (überlappendes Fragment) als sogenannte *silent peaks* konnten mithilfe des R-*Packages* vorgenommen werden. In Abbildung 14 ist ein Beispiel einer *in silico* Assay-Vorhersage mittels R-*Package* MassArray mit Angaben zu analysierbaren und nicht-analysierbaren CpG-Loci

dargestellt. Entsprechend dieser *in silico* Assay-Vorhersage wurden Assays mit prädiktiver Analysierbarkeit bestimmter CpGs für die spätere DNA-Methylierungsanalyse ausgewählt. Die mithilfe der *in silico* Assay-Vorhersage ermittelten Konversionskontrollen, d.h. Fragmente mit einem Cytosin in nicht-CpG-Kontext, konnten im Rahmen der Auswertung der MALDI-TOF-MS-Analyse analysiert werden. Bei effizienter Bisulfit-Konversion wurde hierbei kein Massen-*Peak* erwartet.

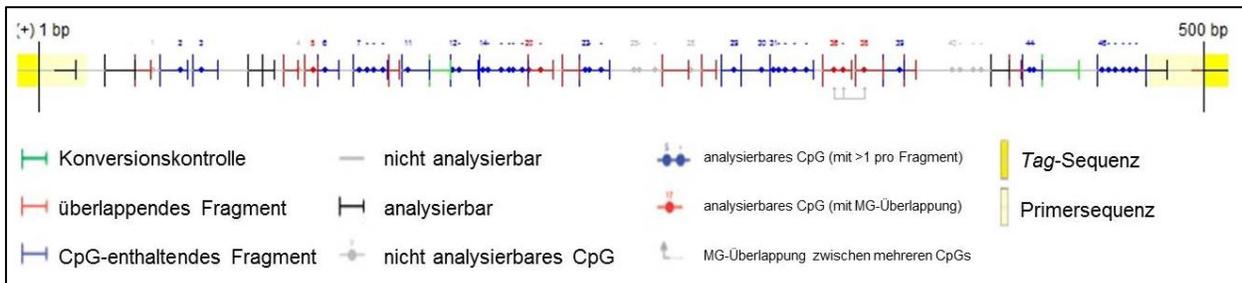


Abbildung 14 Beispiel einer *in silico* Assay-Vorhersage mittels *R-Package MassArray*. Dargestellt ist das prädiktive Fragmentmuster nach *in silico* Fragmentierung (basiert auf einer basenspezifischen Spaltung, d.h. der *homogeneous MassCLEAVE™* (hMC)-Reaktion (Sequenom Inc.) im Rahmen der Epityper-Analyse) anhand eines Beispiel-Assays. Konversionskontrolle: Fragment mit Cytosin in nicht-CpG-Kontext; überlappendes Fragment: Fragmente mit derselben Masse; MG: Molekulargewicht; *Tag*-Sequenz: der Rückwärts-Primer enthält einen T7-Promotor-*Tag* und der Vorwärts-Primer eine 10 bp-*Tag*-Sequenz.

2.2.5.1.2 Nicht-*R-Package*-basierte statistische Analysen mittels *R*-Software

Zur Qualitätskontrolle der 450K-Array-Daten wurde unter Verwendung der *R*-Software Korrelationsmatrices nach Pearson-Methode erstellt. Die AVG-Beta-Werte der 450K-Rohdaten (IDAT-Dateien) wurden hierzu mittels Genome Studio (GS)-Software extrahiert und eine SNP-Korrelation sowie eine CpG-Gesamt-Korrelation²⁵ der AVG-Beta-Werte nach Pearson-Methode berechnet. Die resultierenden Korrelationskoeffizienten-Werte (r) wurden in einer Exceltabelle farblich bedingt formatiert (Wert 0 = blau, 0,97 = gelb, 1 = rot (siehe 3.2.2, Abschnitt „Qualitätskontrolle“ pro Teilprojekt)). Die jeweiligen *R*-Befehle werden anhand eines Beispiels in Abschnitt 8 (Tabelle 36) erläutert. Des Weiteren wurden für verschiedene Analysen, wie BSPPS, qPCR und 450K-Array, die *R*-Befehle `merge()` zum kontrollierten Zusammenführen von Datensätzen, `boxplot()` für Boxplots, `hist()` für Histogramm-Plots, `plot()` für Scatterplots und `abline(lm())` für eine lineare Regression verwendet. Auch *wT*-Tests und Wilcoxon *Signed-Rank* Test (siehe 2.2.4) wurden mittels *R*-Software berechnet.

²⁵ Begriff schließt ch-Loci mit ein

2.2.5.2 Bioinformatische Auswertung mittels Qlucore-Software

Des Weiteren wurde für die die statistischen Analysen der 450K-DNA-Methylierungsanalysen die Qlucore-Software (Qlucore Omics Explorer Version 3.0) angewendet. Eine Importtabelle basierend auf der Normal-GS-Tabelle (Details siehe 2.2.2.2) wurde erstellt, in der zusätzliche Informationen, wie z.B: Diagnose, Alter, oder Gewebetyp, den Proben zugeordnet wurden. Nach dem Einladen der so vorbereiteten Tabellen wurden Hauptkomponentenanalysen (*Principal Component Analysis*, PCA) und hierarchische Clusteranalysen zur Erstellung von sogenannten Heatmaps durchgeführt. In Tabelle 19 sind die statistischen Parameter und die gewählte Normalisierung zur graphischen Darstellung aufgeführt.

Tabelle 19 Statistische Parameter und Normalisierung zur graphischen Darstellung der statistischen Analysen

Analyseform	Statistische Parameter	Normalisierung zur graphischen Darstellung (PCA)	Normalisierung zur graphischen Darstellung (Heatmap)
unsupervidiert	$\sigma/\sigma_{\max} \geq 0,4$	$m = 0$	$m = 0$
ANOVA	$\sigma/\sigma_{\max} \geq 0,4$ $q \leq 1 \times 10^{-6}$	$m = 0$	keine
<i>Student's</i> T-Test	$\sigma/\sigma_{\max} \geq 0,4$ $q \leq 1 \times 10^{-6}$	$m = 0$	keine
Rangregression	$\sigma/\sigma_{\max} \geq 0,2$ $q \leq 3 \times 10^{-4}$	$m = 0$	$m = 0$

σ : Varianz.; q : FDR; m (*Mean*) = 0: Mittelwert aller AVG-Beta-Werte eines Locus werden gleich 0 („*mean-centring*“) gesetzt; keine: keine Normalisierung zur graphischen Darstellung, Einfärbung der AVG-Beta-Werte entspricht dem Absolutwert.

Grundsätzlich wurde zunächst eine unsupervidierte Analyse durchgeführt, um signifikante Segregationen von verschiedenen Gruppen feststellen zu können. Die Normalisierung zur graphischen Darstellung in der Heatmap einer unsupervidierten Analyse entsprach m (*Mean*) = 0 (Mittelwerte aller AVG-Beta-Werte eines Locus sind gleich 0 gesetzt), um so auch geringe Methylierungsunterschiede sichtbar zu machen. Die PCA wurde bei allen Analyseformen mit $m = 0$ dargestellt. Als Varianzfilter wurde $\sigma/\sigma_{\max} \geq 0,4$ für alle Analysen gewählt, wenn dies nicht anders beschrieben wurde. Loci, die eine extreme Varianz der AVG-Beta-Werte über die verschiedenen Proben und somit eine starke Heterogenität zeigten, wurden so ausgeschlossen. Basierend auf diesem so gefilterten Datensatz wurde je nach Fragestellung eine Varianzanalyse (ANOVA) oder ein *Student's* T-Test mit $q \leq 1 \times 10^{-6}$ durchgeführt. In der Heatmap erfolgte hierbei keine Normalisierung zur graphischen Darstellung, sodass eine Einfärbung nach absoluten AVG-Beta-Werten vorlag. Eine Rangregressions-Analyse wurde in dieser Arbeit von Gewebealter-Daten durchgeführt, wobei in der Heatmap-Darstellung die Proben nach Alter sortiert und nicht hierarchisch geclustert dargestellt wurden. Die Normalisierung zur graphischen Darstellung in der Heatmap bei der Rangregressions-Analyse entsprach $m = 0$.

2.2.5.3 Funktionelle Anreicherungsanalysen und Gennetzwerk-Analysen

Funktionelle Anreicherungsanalysen wurden mittels WEB-based GEne SeT AnaLysis Toolkit (WebGestalt) durchgeführt. Für die 450K-Array-Analysen wurden aus der erstellten Hintergrundliste (siehe 2.2.2.2) die Gennamen (UCSC_Refgene_Name) der Loci extrahiert und als Textdatei gespeichert. Bei der WES-Analyse wurde als Hintergrundliste die Annotationsliste der *Enrichment-Library*, hier die SeqCap_EZ_Exome_v2_annotations.txt-Datei, verwendet. Die jeweilige Hintergrundliste und die zu analysierende Liste wurden über die WebGestalt-Benutzeroberfläche importiert. Zur funktionellen Anreicherungsanalyse fanden folgende Parameter (Tabelle 20) Anwendung.

Tabelle 20 Parameter für WEB-based GEne SeT AnaLysis Toolkit (WebGestalt)

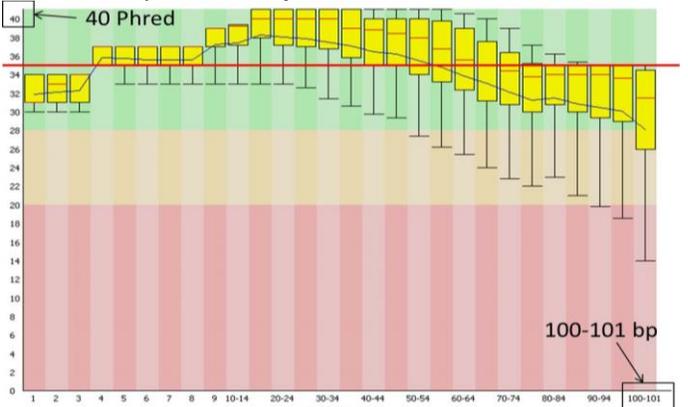
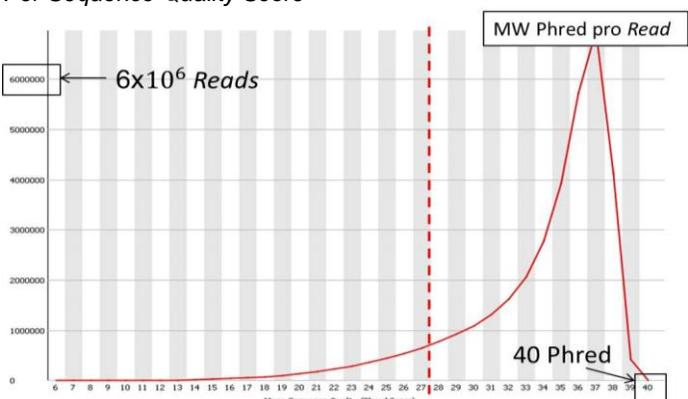
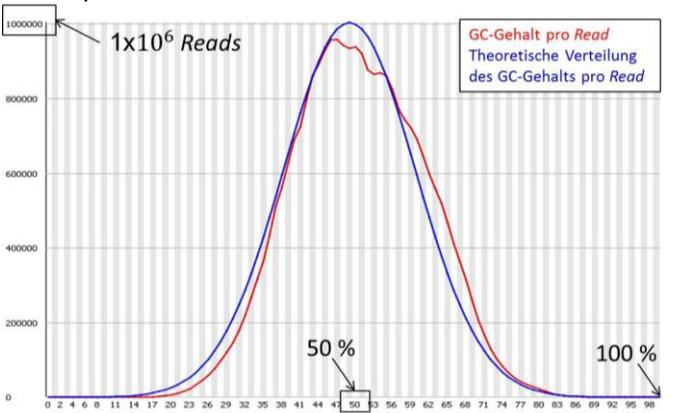
Parameter	Einstellung
ID Typ	hsapiens_gene_symbol
Statistik	hypergeometrisch
Signifikanzlevel	$q \leq 0,001$
<i>Multiple Test Adjustment</i>	nach Benjamini und Hochberg (1995) [297]

Mithilfe der WebGestalt-Oberfläche wurden die eingeladenen Genlisten hinsichtlich der *Gene Ontology* (GO) [299], eine biomedizinische Ontologie, die die drei Bereiche „Zelluläre Komponente“, „Biologischer Prozess“ und „Molekulare Funktion“ umschließt, analysiert. Das Ergebnis der Anreicherungsanalysen wurde mit einem q-Wert angegeben. Pro Anreicherungsanalyse werden im Ergebnisteil dieser Arbeit jeweils die drei am höchsten signifikanten Ergebnisse dargestellt.

Gennetzwerk-Analysen wurden mittels GeneMANIA Gennetzwerk-Analyse-Tool [263] oder STRING-Tool [274] zur Analyse von Protein-Protein Interaktions-Netzwerken durchgeführt (Details zu Software-Tools siehe 2.1.3.2). Die STRING-Analyse wurde unter folgenden Einstellungen durchgeführt: höchste Konfidenz (0,9), keine Text-basierte Assoziation, nicht-verbundene Netzwerkpunkte werden nicht angezeigt, Minimum von fünf Interaktionen.

2.2.5.4 FastQC

Rohdaten der RNA-Sequenzierung wurden mittels FastQC-Software (Version 0.11.2, Babraham Bioinformatics) zur Qualitätskontrolle analysiert. Nachdem die Dateien importiert wurden, erfolgte die Analyse durch die Software und die Resultate wurden als Text- und HTML-Dateien exportiert. Im Folgenden werden die analysierten Qualitätsparameter anhand der Probe HLHS_116 kurz erläutert.

Parameter	Erklärung und Ausschlusskriterien
<p>Per Base Sequence Quality</p> 	<p>Boxplot-Darstellung der Phred-Qualitätswerte (y-Achse), die die Genauigkeit des <i>Base-Callings</i> wiedergeben, d.h. der Wahrscheinlichkeit, dass eine Base durch die Sequenzierungstechnik falsch zugeordnet wurde. Ein Phred-Wert von 30 entspricht einer Wahrscheinlichkeit von einer inkorrekten Base in 1000 Basen (<i>Base-Calling</i>-Genauigkeit = 99,9 %). Die Hintergrundfärbung in grün stellt gute Phred-Werte dar, gelb = mittelmäßig und rot = schlechte Qualität. Die x-Achse stellt die bp-Position der <i>Reads</i> (hier 101 bp lang) dar. Der Median ist als rote Linie in den Boxplots dargestellt, die blaue Linie zeigt den Mittelwert. Als Grenzwert wurde ein Phred-Wert-Median ≥ 25 definiert (horizontale rote Linie). Proben mit einem Phred-Wert-Median < 25 würden aus der Analyse ausgeschlossen werden.</p>
<p>Per Sequence Quality Score</p> 	<p>Mit Hilfe der <i>per Sequence Quality Scores</i> (Phred-Qualitätswerte pro Sequenz) kann erkannt werden, ob ein größerer Teil der Sequenzen einen geringen Phred-Qualitätswert hat. Dies wäre z.B. durch einen weiteren <i>Peak</i> bei geringeren Phred-Werten zu sehen, die auf der x-Achse dargestellt sind. Die y-Achse stellt die Gesamtanzahl aller <i>Reads</i> dar. Als Grenzwert wurde hier ein mittlerer (MW) Phred-Wert pro <i>Read</i> von ≥ 27 definiert (gestrichelte Linie), nach Angaben des Herstellers von FastQC (Babraham Bioinformatics). Dieser Grenzwert entspricht einer Fehlerrate von 0,2 %.</p>
<p>Per Sequence GC Content</p> 	<p>Der <i>per Sequence GC Content</i> (GC-Gehalt pro Sequenz) stellt den GC-Gehalt (Anteil von G- und C-Basen in [%] = $(G+C)/((G+C+A+T) \times 100)$) über jeden Sequenz-<i>Read</i> dar (rote Kurve) und vergleicht diesen mit einer modellierten Normalverteilung der GC-Werte (theoretische Verteilungskurve in blau). Eine theoretische Verteilung des GC-Gehalts des zugrundeliegenden Transkriptoms wird durch die FastQC-Software anhand der beobachteten Daten berechnet. Die x-Achse stellt den mittleren GC-Gehalt [%] und die y-Achse die Gesamtanzahl aller <i>Reads</i> dar. Eine veränderte Form der roten Kurve gegenüber der theoretischen Verteilungskurve kann z.B. auf Kontaminationen in der <i>Library</i> hinweisen. Wenn mehr als 15 % der <i>Reads</i> eine Abweichung von der theoretischen Verteilung zeigten, wurden diese aus der Analyse ausgeschlossen.</p>

3 Ergebnisse

Das Ziel dieser Arbeit war die Analyse ursächlicher Zusammenhänge in der Entstehung des hypoplastischen Linksherzsyndroms (HLHS), einer kongenitalen Herzfehlbildung (CHD). Bis heute ließen sich nur wenige Gene in einer möglichen Assoziation mit dem HLHS identifizieren [106, 107, 221, 196, 222, 225, 226], wobei der bisherige Fokus auf molekulargenetischen Untersuchungen der Keimbahn lag. Potenziell können jedoch auch auf somatischer Ebene, d.h. im betroffenen Herzgewebe, Ereignisse auftreten, die in der Pathogenese des HLHS eine Rolle spielen. Um daher auf den verschiedenen Ebenen die molekularen Mechanismen in der Entstehung des HLHS, zu untersuchen, wurden in dieser Arbeit folgende Analysen durchgeführt (Abbildung 15):

- (1) Analysen auf **Keimbahnebene** (siehe 3.1), welche genetische Analysen hinsichtlich syndrom-assoziiierter CHDs und nicht-syndromaler HLHS-Fälle umfassten
- (2) Analysen auf **somatischer Ebene** von Herzgewebeproben (siehe 3.2). Diese beinhalteten Gesamt-Exom-Analysen, epigenetische Analysen in Form von DNA-Methylierungsanalysen, sowie Transkriptom-Analysen.
- (3) Abschließend wurden in einer **integrativen Analyse** (3.3) Genom, Epigenom und Transkriptom untersucht.

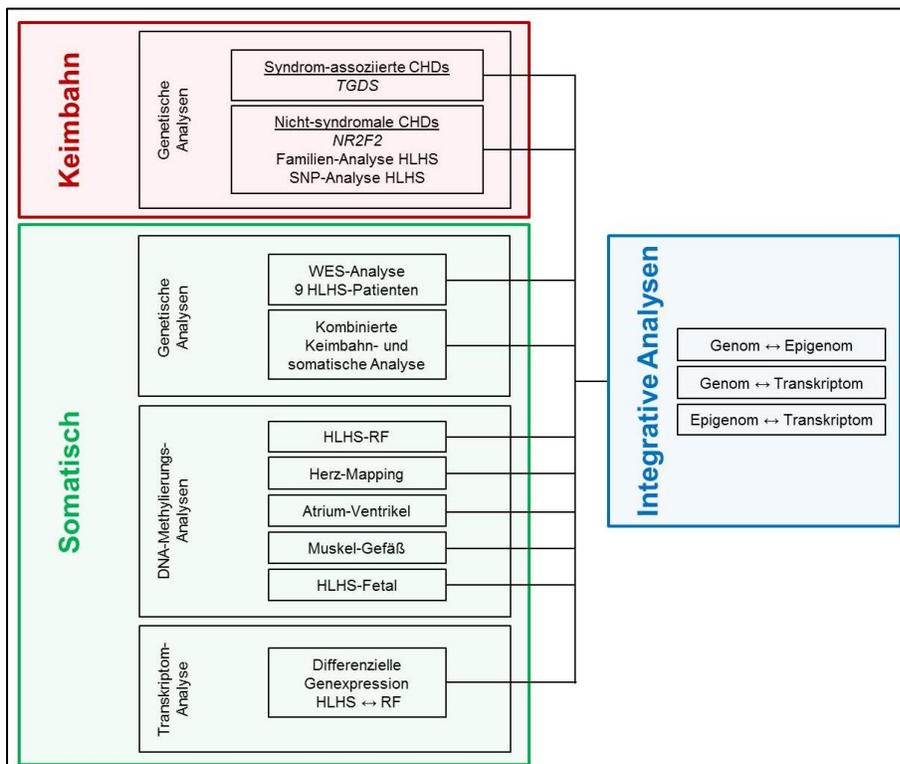


Abbildung 15 Schematische Übersichtsdarstellung der Analysen und jeweiligen Teilprojekte auf Keimbahn- und somatischer Ebene, sowie integrativer Analysen. WES: Gesamt-Exom-Sequenzierung; RF: Rechtsherzfehlbildungen; Fetal: fetales Herzgewebe gesunder Individuen

3.1 Analysen auf Keimbahnebene

Der Einfluss einer genetischen Komponente bei Fällen mit CHDs wurde erstmals in 1982 vermutet [300] und durch weitere epidemiologische Daten [301, 302] und Familienstudien bestätigt [105, 110, 234]. Um daher auf Ebene der Keimbahn ursächliche Zusammenhänge zu untersuchen, wurden in der vorliegenden Arbeit genetische Analysen eines syndrom-assoziierten CHDs und von nicht-syndromalen HLHS-Fällen durchgeführt.

3.1.1 Identifizierung von Mutationen in *TGDS* als Beispiel einer syndrom-assoziierten kongenitalen Herzfehlbildung

Um mögliche pathogene Ereignisse auf Keimbahnebene von syndrom-assoziierten CHD-Fällen zu analysieren, wurden in dieser Arbeit als ein Beitrag zur Studie von Ehmke et al. (2014) [104] Gesamt-Exom-Sequenzierungen (WES, siehe 2.2.1.7) von drei Familien mit Catel-Manzke-Syndrom (Familie 1 – 3, siehe 2.1.1 Tabelle 2) durchgeführt. Patienten mit diesem Syndrom haben in 40 % der Fälle eine Herzfehlbildung [103]. Als kausales Gen für das beobachtete autosomal-rezessive Vererbungsmodell des Catel-Manzke-Syndroms konnte in der vorliegenden Arbeit das *TGDS*-Gen identifiziert werden [104]. Die mittels WES detektierten Mutationen in *TGDS* der drei Familien 1 – 3 konnten anschließend durch Sanger-Sequenzierung (siehe 2.2.1.6) verifiziert werden.

Folgende Genvarianten in *TGDS* (NM_014305.2) konnten in den drei analysierten Familien (Familie 1 – 3, siehe 2.1.1 Tabelle 2, Stammbaum in Abbildung 16 A) mittels WES detektiert und anschließend verifiziert werden: c.298G>T (p.Ala100Ser) und c.294T>G (p.Phe98Leu). Eine homozygote Ausprägung der Variante c.298G>T ließ sich bei zwei Indexpatienten, je von Familie 1 und 2 nachweisen (Tabelle 21). Bei Familie 3 konnte eine biparentale Vererbung der *compound* heterozygoten Mutation c.298G>T/c.294T>G beim Indexpatient 3 detektiert werden. Dieser Patient wies mit einem Ventrikelseptumdefekt (VSD) einen kardialen Phänotyp und somit einen syndrom-assoziierten CHD auf (Tabelle 21).

Tabelle 21 Mittels WES nachgewiesene Mutationen in *TGDS* (NM_014305.2) bei Familien mit Catel-Manzke-Syndrom

Proben-ID	Familien-Zuordnung	Diagnose	kardiale Diagnose	Position CHR (GRCh37/hg19)	Mutation (Konseq.)	Typ	PPH
D12/579	Familie_1_Index	CMS	-	chr13:95243122	c.298G>T (p.Ala100Ser)	homozygot	p. d.
D12/581	Familie_1_Sohn	normal	-	chr13:95243122	c.298G>T (p.Ala100Ser)	heterozygot	p. d.
D12/775	Familie_2_Index	CMS	-	chr13:95243122	c.298G>T (p.Ala100Ser)	homozygot	p. d.
D12/776	Familie_2_Mutter	normal	-	chr13:95243122	c.298G>T (p.Ala100Ser)	heterozygot	p. d.
D11/527	Familie_3_Index	CMS	VSD*	chr13:95243122	c.298G>T (p.Ala100Ser)	compound heterozygot	p. d.
				chr13:95243126	c.294T>G (p.Phe98Leu)		-
D11/528	Familie_3_Mutter	normal	-	chr13:95243122	c.298G>T (p.Ala100Ser)	heterozygot	p. d.
D11/1087	Familie_3_Vater	normal	-	chr13:95243126	c.294T>G (p.Phe98Leu)	heterozygot	-

Familien-Zuordnung nach 2.1.1.1 (Tabelle 2); *VSD mit spontanem Verschluss; Position CHR: chromosomale Position (nach GRCh37/hg19); Mutation (Konseq.): Nomenklatur Mutation in kodierender Sequenz und Konsequenz auf Proteinebene in Klammern; PPH: Polyphen-2; CMS: Catel-Manzke-Syndrom; p.d.: *probably damaging*

Bei der Genvariante c.298G>T (p.Ala100Ser) handelt es sich um eine bekannte Variante (rs140430952), welche heterozygot in 9 von 4291 Individuen nach EVS-Datenbank und in 34 von 66646 Individuen nach ExAC-Datenbank vorliegt. Eine homozygote Ausprägung der Mutation lag dagegen in keiner der Datenbanken vor. Nach prädiktiver Mutationskonsequenz (Polyphen-2 (Polymorphism Phenotyping v2; [272])) wurde die Mutation p.Ala100Ser als *probably damaging* eingestuft. Die Genvariante c.294T>G (p.Phe98Leu) wurde bisher noch in keiner Datenbank annotiert und stellt somit eine neue, unbekanntere Variante dar.

Mittels Sanger-Sequenzierung (siehe 2.2.1.6) konnten die in Tabelle 21 aufgeführten Mutationen in *TGDS* bei Indexpatienten mit Catel-Manzke-Syndrom und zugehörigen Familienmitgliedern verifiziert werden. Die Stammbäume der drei analysierten Familien mit Angaben zum *TGDS*-Genotyp und die entsprechenden Elektropherogramme sind in Abbildung 16 dargestellt.

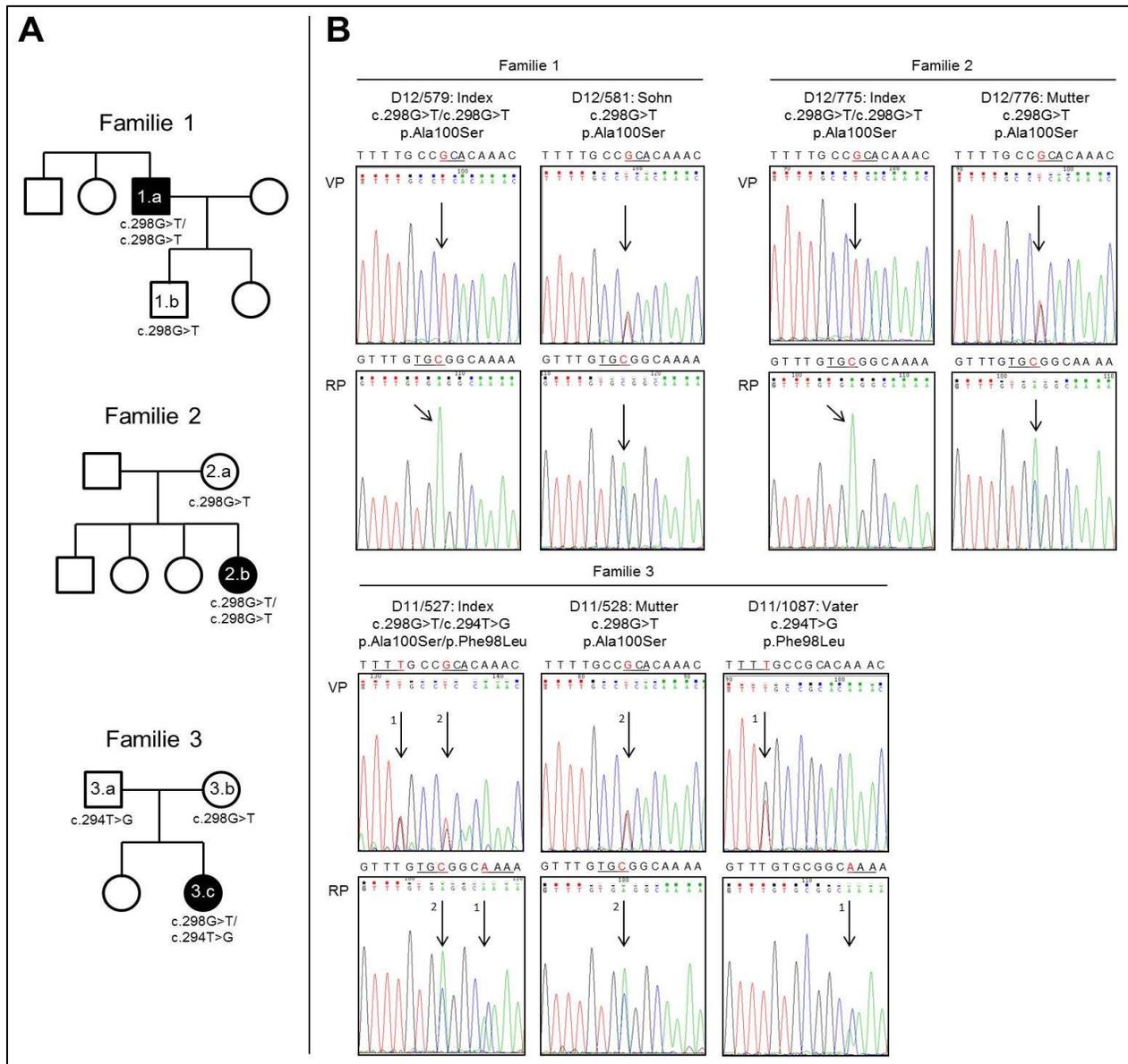


Abbildung 16 Übersicht der nachgewiesenen Mutationen in *TGDS* (NM_014305.2) in drei analysierten Familien mit Catel-Manzke-Syndrom. A: Stammbäume und *TGDS*-Genotypen der drei analysierten Familien mit Catel-Manzke-Syndrom. Folgende DNA-Proben der Familienmitglieder wurden sequenziert (Tabelle 21): 1.a: D12/579, 1.b: D12/581, 2.a: D12/776, 2.b: D12/775, 3.a: D11/1087, 3.b: D11/528, 3.c: D11/527. Eine schwarze Füllung symbolisiert einen Indexpatienten mit Catel-Manzke-Syndrom, während eine weiße Füllung für nicht betroffene Familienmitglieder steht. Mutationen in *TGDS* sind unterhalb der jeweils analysierten Familienmitglieder angegeben. (modifiziert nach Ehmke et al., (2014) [104]). B: Dargestellt sind die Elektropherogramme mit den *TGDS*-Mutationen (siehe Pfeile) der analysierten Indexpatienten und zugehörigen Familienmitgliedern von Familie 1, 2 und 3 (Tabelle 21). Pro Familie ist oben jeweils die Sequenz des Vorwärts-Primers (VP) und unten die des Rückwärts-Primers (RP) dargestellt. Rote Buchstaben stellen die mutierte Base dar, unterstrichene Buchstaben die mutierte Kodonsequenz der Aminosäuren. Pfeil 1: c.298G>T; Pfeil 2: c.294T>G.

Zusammenfassend konnten somit homozygote und *compound* heterozygote Mutationen in *TGDS* als Beispiel eines syndrom-assoziierten CHDs in drei Familien mit Catel-Manzke-Syndrom als kausale Genvarianten identifiziert und verifiziert werden.

3.1.2 Analysen nicht-syndromaler Herzfehlbildungen

Zur Ätiologie isolierter, nicht-syndromaler CHD-Fälle (CHDs ohne zusätzliche extrakardialen oder fazialen Auffälligkeiten), die einen Großteil der CHDs darstellen, ist bisher nur wenig bekannt [69, 77]. Daher erfolgte im Rahmen dieser Arbeit eine Beteiligung an einer Studie von Al Turki et al. (2014) [303], welche zeigen konnte, dass Genvarianten in *NR2F2* in einem ursächlichen Zusammenhang mit (nicht-syndromalen) atrioventrikulären Septumdefekten (AVSDs) und LVOTO-Phänotypen (linksventrikuläre Ausflusstrakt-Obstruktionen) stehen. Außerdem wurde in dieser Arbeit im Rahmen einer WES-Studie zu Familien mit LVOTOs (Hitz und Hoff et al., Publikation in Vorbereitung, siehe 11) eine Familie mit zwei nicht-syndromalen HLHS-Fällen hinsichtlich krankheitsrelevanter Genvarianten untersucht. Des Weiteren erfolgte mittels SNP-Array (Genome-Wide Human SNP Array 6.0) eine Analyse einer Familie mit einem nicht-syndromalen HLHS-Fall, um krankheitsassoziierte CNVs zu identifizieren.

3.1.2.1 Identifizierung einer *de novo* Mutation in *NR2F2* bei einem nicht-syndromalen HLHS-Fall

In der Studie von Al Turki et al. (2014) [303] konnten kausale Genvarianten in *NR2F2* (NM_021005.3) bei Familien und Eltern-Index-Trios mit isoliert auftretenden AVSDs und LVOTO-Phänotypen, wie auch dem HLHS, durch WES detektiert werden. In der vorliegenden Arbeit konnte als Beitrag zu dieser Studie von Al Turki et al. (2014) eine *de novo* Splice-Donor-Mutation in *NR2F2* in einem Eltern-Index-Trio eines Patienten mit HLHS verifiziert werden (Familie 7 siehe 2.1.1.1 Tabelle 2).

Die Mutation c.970+1G>A in *NR2F2*, welche in der Donor-Splice-Site am 5'-Ende des Introns liegt, konnte mittels Sanger-Sequenzierung (siehe 2.2.1.6) nachgewiesen werden. In Abbildung 17 A ist die Splice-Donor-Mutation im Kontext mit dem Exon-Intron-Übergang im Elektropherogramm dargestellt.

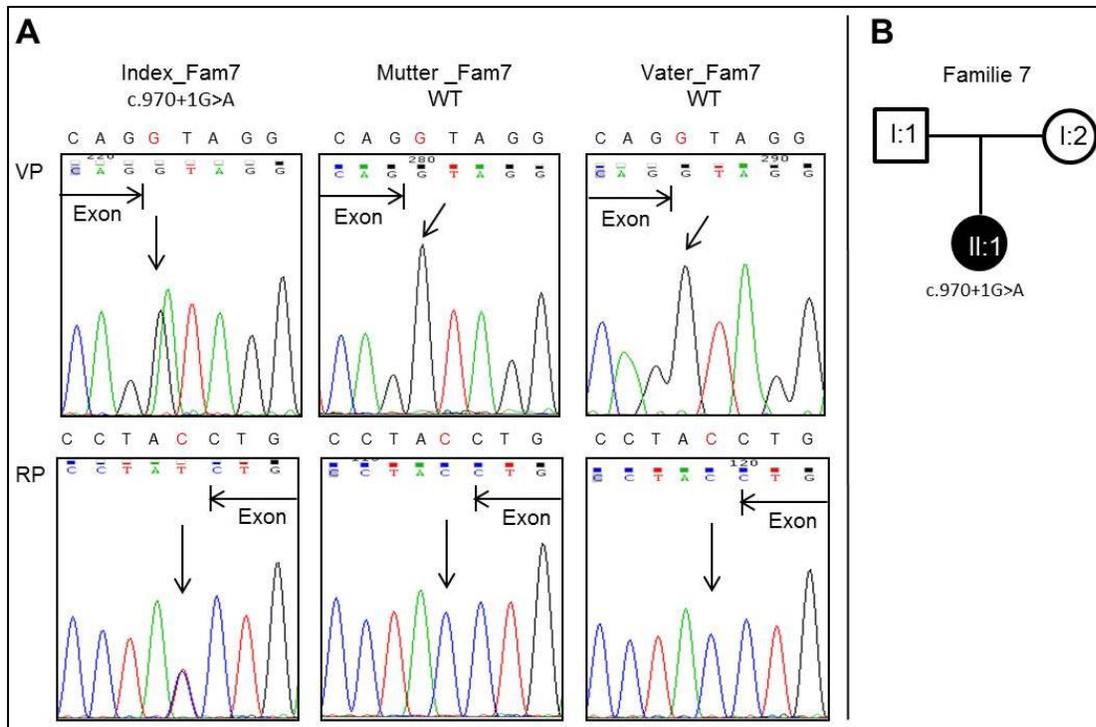


Abbildung 17 Übersicht der mittels Sanger-Sequenzierung nachgewiesenen *de novo* Splice-Donor-Mutation c.970+1G>A in *NR2F2* (NM 021005.3) eines Eltern-Index-Trios (Familie 7) eines HLHS-Patienten. A: Dargestellt sind die Elektropherogramme von HLHS-Indexpatient (mit der *de novo* Mutation c.970+1G>A), Mutter (Wildtyp (WT)-Sequenz) und Vater (WT-Sequenz) mit der Position der *NR2F2*-Mutation als Pfeil markiert. Das obere Elektropherogramm zeigt die Sequenz des Vorwärts-Primers (VP) und das untere die des Rückwärts-Primers (RP). Rote Buchstaben stellen die mutierte Base dar. Die Exon-Regionen sind durch Pfeile markiert. B: Stammbaum der Familie 7 (modifiziert nach Al Turki et al., (2014) [303]. Die Eltern I:1 (Vater) und I:2 (Mutter) sind phänotypisch gesund (weiße Füllung) und sie wiesen keine Mutation in *NR2F2* auf. Beim Indexpatienten mit HLHS (II:1) ist der Genotyp der Splice-Donor-Mutation c.970+1G>A in *NR2F2* angegeben.

Diese Genvariante c.970+1G>A ließ sich beim Indexpatient mit HLHS in heterozygoter Ausprägung mittels Sanger-Sequenzierung nachweisen, während die phänotypisch unauffälligen Eltern diese Mutation nicht trugen (Abbildung 17 A und B). Somit handelte es sich um eine *de novo* Mutation. Diese *de novo* Splice-Donor-Mutation in *NR2F2*, kodierend für den liganden-induzierbaren COUP-II Transkriptionsfaktor²⁶, stellt eine bisher unbekannte Genvariante dar. Prädiktiv führt diese Mutation zu einem trunkierten Protein [303].

Zusammenfassend ließ sich eine kausale, bisher unbekannte *de novo* Splice-Donor-Mutation in *NR2F2* bei einem Patienten mit HLHS als Beispiel eines nicht-syndromalen CHDs nachweisen.

²⁶ Informationen zu vollständigen Gennamen bzw. Informationen über die Funktion des Genprodukts grundsätzlicher Natur wurden entsprechend der Angaben der Online-Datenbank GeneCards® [262] (siehe 2.1.3.2 Tabelle 12) übernommen und im nachfolgenden Text daher nicht mit einer gesonderten Referenz versehen.

3.1.2.2 Analyse einer Familie mit zwei HLHS-Fällen: Identifizierung von *C15orf62*, *TNKS1BP1* und *PRF1* als potenziell krankheitsrelevante Kandidatengene nach autosomal dominantem Vererbungsmodell mit reduzierter Penetranz

Bei Familien mit rekurrent auftretenden CHD-Fällen kann häufig eine reduzierte Penetranz detektierter Mutationen beobachtet werden [78, 236, 246, 304]. Dieses autosomal dominante Vererbungsmodell mit reduzierter Penetranz konnte in der vorliegenden Arbeit im Rahmen einer WES-Studie zu Familien mit LVOTOs (Hitz und Hoff et al., Publikation in Vorbereitung (siehe 11)) bei drei potenziell krankheitsrelevanten Genen *C15orf62* (NM_001130448.2), *TNKS1BP1* (NM_033396.2) und *PRF1* (NM_005041.4) in einer Familie mit zwei HLHS-Fällen nachgewiesen werden.

Zur Familie 5 (siehe 2.2.1.7 Tabelle 13, Stammbaum Abbildung 18 A) stand von drei Familienmitgliedern, dem Vater und jeweils den beiden Indexpatienten mit HLHS (HLHS_130_Index-1, und -2), ein analysierter WES-Datensatz²⁷ zur Verfügung. Im Abgleich mit den SNP-*Calling*-Daten der RNA-Sequenzierung von HLHS_130_Index-1 ließen sich alle mittels WES-Methode identifizierten Einzelnukleotidvarianten (SNVs) auch beim SNP-*Calling* identifizieren, sodass von einem validen Datensatz ausgegangen werden konnte. Anhand eines Priorisierungs-*Workflows* [Hitz, 283] wurde der WES-Datensatz zuvor hinsichtlich seltener dominanter (ExAC AF < 0,1 %) und rezessiver²⁸ bzw. X-chromosomal-vererbter (ExAC AF < 1 %) Genvarianten nach dem wahrscheinlichsten Vererbungsmodell im Kontext bekannter Diagnosen des Familienstammbaums (Abbildung 18 A) gefiltert. Nach klassischem Mendelschen Vererbungsmodell²⁹ [305] ließen sich in dieser Familie keine Kandidatengene identifizieren, sodass entsprechend des Priorisierungs-*Workflows* hinsichtlich Kandidatengenen nach dem Vererbungsmodell „autosomal dominant mit reduzierter Penetranz“ analysiert wurde. Eine Vererbung nach letztgenanntem Modell wurde über die Generationen mütterlicherseits vermutet, da in diesen Generationen Totgeburten auftraten (Abbildung 18 A, II:5 und III:3). Zwar war die Todesursache ungeklärt, jedoch zählen CHDs zu den häufigsten Ursachen bei Totgeburten [50], wobei zyanotische CHDs (wozu das HLHS zählt) hierbei als häufigste Todesursache beschrieben wurden [306]. Zusätzlich zu den mittels WES analysierten drei Proben (Vater, Indexpatient-1 und -2) wurde über die Biomaterialbank des Kompetenznetzes

²⁷ Die Analyse der WES-Daten erfolgte im Vorfeld durch den Kollegen Dr. med Marc-Phillip Hitz.

²⁸ homozygot oder *compound* heterozygot bei Indexpatienten vorliegend

²⁹ einfacher Erbgang: Vererbungsvorgang von Merkmalen, deren Ausprägung von nur einem Gen bestimmt wird

Angeborene Herzfehler e.V. DNA zweier weiterer Familienmitglieder (Mutter und gesundes Geschwisterkind (HLHS_130_Bruder)), zur Verfügung gestellt. Es erfolgten in dieser Arbeit Sanger-Sequenzierungsanalysen von 21 Kandidatengenen in allen fünf Familienmitgliedern (Primer siehe Abschnitt 8 Tabelle 31). Die Ergebnisse der Analyse sind in Abschnitt 8 in Tabelle 37 zusammengefasst.

Die Familien-Analyse mittels Sanger-Sequenzierungsanalysen resultierte in drei potenziell krankheitsrelevanten Genen: *C15orf62* (NM_001130448), *TNKS1BP1* (NM_033396) und *PRF1* (NM_005041). In diesen Genen konnten heterozygote Mutation in beiden Indexpatienten mit HLHS (Abbildung 18 A, IV:1 und IV:1) sowie bei der Mutter (Abbildung 18 A, III:2) nachgewiesen werden. Die phänotypisch gesunden Familienmitglieder – Vater und drittes Geschwisterkind (Abbildung 18 A, III:1 und IV:3) – wiesen keine Mutation auf. Entsprechend des hypothetischen Vererbungsmodells lag eine reduzierte Penetranz vor, d.h. die Mutter zeigte dieselben Mutationen wie die betroffenen Indexpatienten ohne eine phänotypische Ausprägung der Erkrankung.

In *C15orf62* konnte eine bisher unbekannte heterozygote Missense-Mutation (c.335C>G (p.Ser112Trp)) nachgewiesen werden (Abbildung 18 A, B), die nach Polyphen-2 als *possibly damaging* eingestuft wird. In *TNKS1BP1* konnte eine seltene (ExAC-AF < 3x10⁻⁵, 2/66708 Individuen) heterozygote Missense-Mutation (c.4198G>A (p.Gly1400Arg)), nach Polyphen-2 als *probably damaging* eingestuft, nachgewiesen werden (Abbildung 18 A, C). Eine bisher unbekannte heterozygote Stop-Mutation (c.757G>T (p.Glu253*)) konnte in *PRF1* detektiert werden (Abbildung 18 A, D), welche auf Proteinlevel nach ELM-Datenbank (Eukaryotic Linear Motif [261], siehe 2.1.3.2) eine Trunkierung der Hälfte des Proteins bewirken würde.

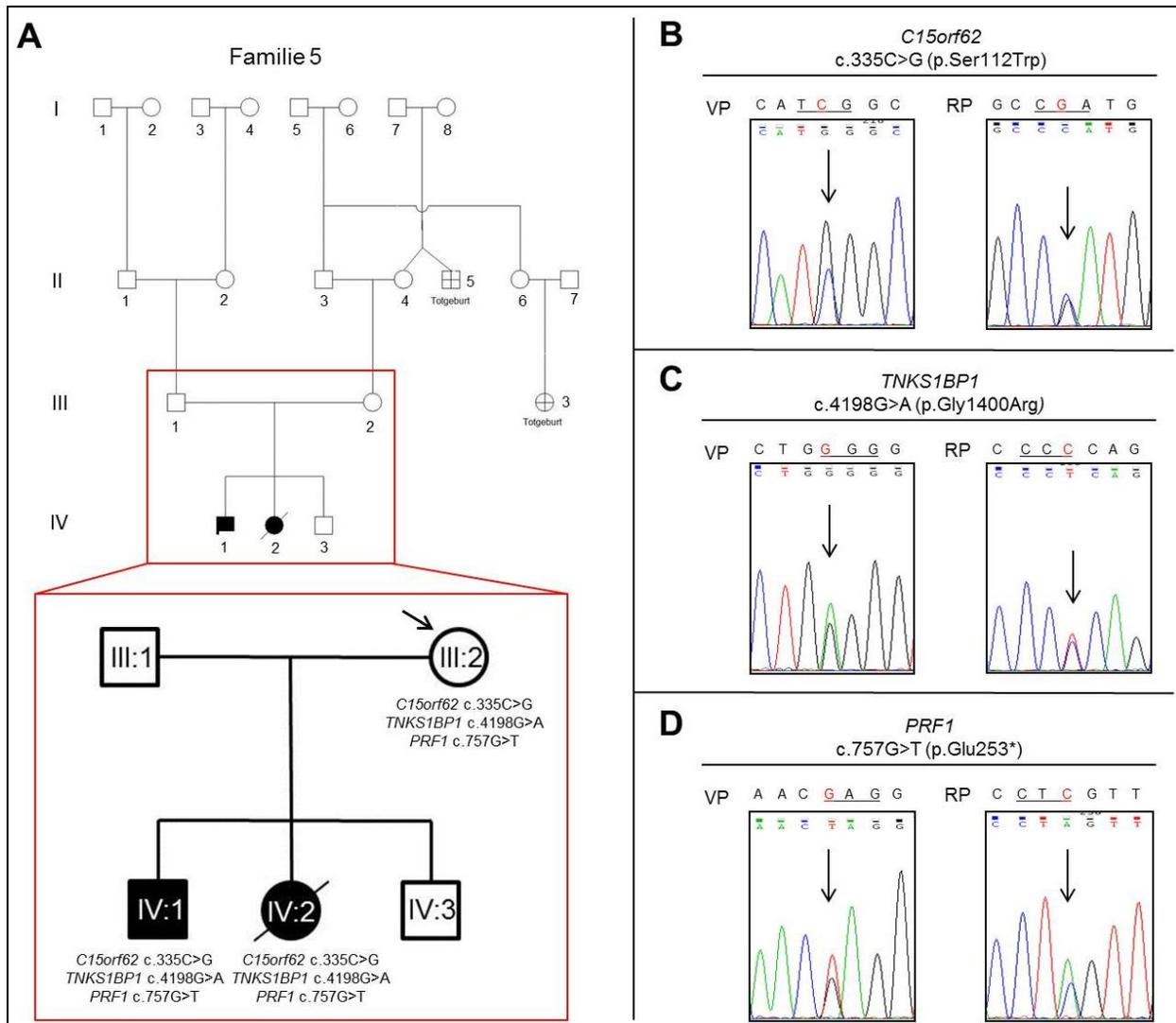


Abbildung 18 Übersicht der in Familie 5 mittels Sanger-Sequenzierung nachgewiesenen Mutationen in *C15orf62* (NM 001130448.2), *TNKS1BP1* (NM 033396.2) und *PRF1* (NM 005041.4). A: Stammbaum der analysierten Familie 5 (siehe 2.1.1.1, Tabelle 2) mit Genotypen der Mutation in *C15orf62* (c.335C>G), *TNKS1BP1* (c.4198G>A) und *PRF1* (c.757G>T) unterhalb der jeweils analysierten Familienmitglieder. Eine schwarze Füllung symbolisiert einen Indexpatienten mit HLHS (durchgestrichenes Symbol = verstorben), während eine weiße Füllung für nicht betroffene Familienmitglieder steht. Der gesunde Vater (III:1, HLHS_130_Vater) und das gesunde Geschwisterkind (IV:3, HLHS_130_Bruder) wiesen keine Mutationen auf. Die phänotypisch gesunde Mutter (III:2, HLHS_130_Mutter) sowie die Geschwister mit HLHS, Indexpatient-2 (IV:1, HLHS_130_Index-2) und Indexpatient-1 (IV:2, HLHS_130_Index-1), zeigten jeweils eine heterozygote Ausprägung der Mutationen in *C15orf62* (c.335C>G), *TNKS1BP1* (c.4198G>A) und *PRF1* (c.757G>T). Der Pfeil an III:6 signalisiert das autosomal dominante Vererbungsmodell reduzierter Penetranz über frühere Generationen mütterlicherseits, in denen Totgeburten (Kreuz-Darstellung der Stammbaum-Symbole) bei II:5 und III:3 auftraten. B: Elektropherogramm der Mutation in *C15orf62* (c.335C>G (p.Ser112Trp)) und C: Elektropherogramm der Mutation in *TNKS1BP1* (c.4198G>A (p.Gly1400Arg)). D: Elektropherogramm der Mutation in *PRF1* (c.757G>T (p.Glu253*)). Die Mutationen in Abbildung B, C und D sind beispielhaft anhand der DNA-Probe HLHS_130_Index-1 (IV:2) dargestellt und durch Pfeile markiert. Oberhalb des Elektropherogramms ist jeweils die Sequenz des Vorwärts-Primers (VP) und des Rückwärts-Primers (RP) dargestellt. Rote Buchstaben stellen die mutierte Base dar, unterstrichene Buchstaben die mutierte Kodonsequenz der Aminosäuren.

Zusammenfassend konnten bisher unbekannte bzw. seltene Missense-Mutationen, als *possibly* bzw. *probably damaging* eingestuft, in *C15orf62* und *TNKS1BP1* und eine bisher unbekannte

Stop-Mutation in *PRF1* nach autosomal dominantem Vererbungsmodell mit reduzierter Penetranz als potenziell krankheitsrelevante Genvarianten in einer Familie mit zwei HLHS-Fällen nachgewiesen werden.

3.1.2.3 SNP-Analyse eines nicht-syndromalen HLHS-Falls im familiären

Kontext

Kopienzahlvariationen (CNVs) liegen als ursächliche pathogene Variationen bei ca. 3 – 10 % der nicht-syndromalen CHD-Fälle zugrunde [87, 91, 93]. Nach Hitz et al. (2012) [89] konnten vererbte, sowie *de novo* CNVs in Assoziation zu linksseitigen (LS)-CHDs, zu denen auch das HLHS zählt, nachgewiesen werden. In dieser Arbeit wurde daher mittels Genome-Wide Human SNP Array 6.0 (SNP-Array, siehe 2.2.1.8) eine Analyse einer Familie mit einem nicht-syndromalen HLHS-Fall durchgeführt, um krankheitsassoziierte CNVs identifizieren zu können.

Die Familie 6 (siehe 2.1.1, Tabelle 2) mit einem nicht-syndromalen HLHS-Fall, HLHS_132_Index (pB-Probe zu HLHS_132), wurde hierzu hinsichtlich *de novo* CNVs, d.h. CNVs die ausschließlich beim Indexpatienten HLHS_132_Index auftraten, untersucht. Des Weiteren wurden Kopienzahlverluste- bzw. Zugewinne in bekannten chromosomalen Regionen, die mit einem human kardialen Phänotyp nach der Studie von Thorsson et al. (2015) [307] assoziiert sind, mittels ChAS-Software (Affymetrix® Chromosome Analysis Suite Software Version 2.1) analysiert (Details zu Filterkriterien siehe 2.2.1.8).

Es konnten keine *de novo* CNVs beim Indexpatienten HLHS_132_Index festgestellt werden. Kosegregierende CNVs, jeweils beim Indexpatient und Vater vorliegend, umfassten (1) heterozygote Verluste in 3p12.3 (0,19 Mb), in 10q21.1 (0,29 Mb), in 12p13.31 (0,1 Mb) und in 19p12 (0,11 Mb), und (2) heterozygote Zugewinne in 16p11.2 (0,28 Mb) und in 17q12 (0,1 Mb) (Tabelle 22). Ein 0,11 Mb-Verlust in 9p11.2 lag heterozygot in beiden Elternteilen und dem Indexpatienten vor. Gene innerhalb der CNV-Region umfassten die Pseudogene *FAM86DP* (3p12.3) und *ZNF826P* (19p12), *LOC*-Gene (uncharakterisierte nicht-kodierende RNAs) in 16p11.2, und in 17q12 Gene kodierend für Chemokinliganden, sowie das Gen *TBC1D3B* (Tabelle 22). Bei allen CNVs handelte es sich um bekannte strukturelle Varianten, die nach DGV-Datenbank bereits in Kontroll-Individuen einer weltweiten Population detektiert werden konnten. Lediglich der 0,29 Mb-Verlust in 10q21.1, welcher zuvor in sieben

Kontroll-Individuen nach DGV-Datenbank detektiert wurde, stellte eine seltene, und damit potenziell krankheitsassoziierte CNV dar, während alle weiteren CNVs als *benign* eingestuft wurden (Tabelle 22: Spalte „Signif.“).

Tabelle 22 CNVs von Indexpatient HLHS_132_Index der mittels SNP-Array analysierten Familie 6

CHR-Bande	Typ	Position CHR (GRCh37/hg19)	Größe [Mb]	Gene in CNV-Region	DGV [Anzahl]	Überlappung CNV [Fall]	Signif.
3p12.3	Verlust	3:75427094-75620883	0,19	<i>FAM86DP</i>	39	Vater	b
9p11.2	Verlust	9:44727846-44837515	0,11	-	33	Vater, Mutter	b
10q21.1	Verlust	10:57369116-57656103	0,29	-	7	Vater	p.k.
12p13.31	Verlust	12:9633857-9735686	0,1	-	25	Vater	b
16p11.2	Zugewinn	16:34470703-34746181	0,28	<i>LOC283914</i> <i>LOC146481</i> <i>LOC10013070</i> <i>0</i>	47	Vater	b
17q12	Zugewinn	17:34439965-34543331	0,1	<i>TBC1D3B</i> <i>CCL3L3</i> <i>CCL3L1</i> <i>CCL4L1</i> <i>CCL4L2</i>	52	Vater	b
19p12	Verlust	19:20603317-20716377	0,11	<i>ZNF826P</i>	38	Vater	b

CHR-Bande: mit CNV überlappende Chromosomen-Bande; Position CHR (GRCh37/hg19): chromosomale Position der CNV nach GRCh37/hg19; Größe [Mb]: Größe CNV in Mb (Megabasenpaare); DGV: Anzahl bekannter Varianten nach DGV-Datenbank überlappend mit detektierten CNVs; Überlappung CNV [Fall]; Angabe zur Detektierbarkeit der CNV in Vater (HLHS_132_Vater) bzw. Mutter (HLHS_132_Mutter) des Indexpatienten (siehe 2.1.1, Tabelle 2, Familie 6); Signif.: Signifikanz (p.k.: potenziell krankheitsassoziiert, b: *benign*).

Es konnten keine Kopienzahlverluste- bzw. Zugewinne in bekannten chromosomalen Regionen, die mit einem human kardialen Phänotyp assoziiert sind [307], detektiert werden. Auch eine „Lockerung“ der Analyseparameter mit Verringerung der minimalen CNV-Länge auf 50 kbp ergab keine *de novo* CNVs, seltene CNVs (mit DGV-Anzahl < 20), noch kardial assoziierte CNVs bei HLHS_132_Index.

Zusammenfassend konnten mittels SNP-Analyse von HLHS_132_Index im familiären Kontext weder *de novo* CNVs, noch CNVs in Assoziation zu CHD-assoziierten Genen bzw. chromosomalen Regionen detektiert werden, sodass keine kausalen CNVs nachweisbar waren. Lediglich eine seltene und damit potenziell krankheitsassoziierte CNV konnte in 10q21.1 (0,29 Mb-Verlust) festgestellt werden. Letztere war jedoch auch im phänotypisch gesunden Vater des Patienten detektierbar.

3.2 Analysen auf somatischer Ebene

In Anbetracht der Tatsache, dass sich bisher nur 15 - 20 % der CHD-Fälle durch genetische Faktoren in der Keimbahn oder teratogene Einflüsse erklären lassen [47, 66, 69, 77], könnten Ereignisse auf somatischer Ebene (im betroffenen Herzgewebe) ein alternatives genetisches Modell der Pathogenese von CHDs darstellen. Daher wurden in der vorliegenden Arbeit Herzgewebeproben von Patienten mit CHDs, insbesondere des HLHS mithilfe verschiedener molekulargenetischer, NGS- und Array-basierter Methoden untersucht. Die Ergebnisse zu den Gesamt-Exom-Analysen, den epigenetischen Analysen in Form von DNA-Methylierungsanalysen, sowie den Transkriptom-Analysen sind in den folgenden Abschnitten aufgeführt.

3.2.1 Genetische Analysen an Herzgewebe von Patienten mit HLHS

Um mögliche pathogene somatische Mutationen im betroffenen Gewebe selbst, hier dem Herzgewebe, zu analysieren, wurden Gesamt-Exom-Sequenzierungen (WES, siehe 2.2.1.7) an Herzgewebe von Patienten mit HLHS durchgeführt. Im folgenden Abschnitt werden die Ergebnisse der WES-Analyse von IAS-Gewebe von neun Patienten mit HLHS dargestellt (3.2.1.1). Des Weiteren wurde eine kombinierte Keimbahn- und somatische Analyse in einer Familie mit HLHS durchgeführt, um mögliche Herzgewebe-spezifische krankheitsrelevante Mutationen zu detektieren (siehe 3.2.1.2).

3.2.1.1 WES-Analyse an IAS-Gewebe von Patienten mit HLHS identifiziert

56 seltene bzw. bisher unbekannte *Loss-Of-Function*-Mutationen

Die Ergebnisse zur WES-Analyse an IAS-Gewebe von neun Patienten mit HLHS – HLHS_4, HLHS_8, HLHS_12, HLHS_13, HLHS_15, HLHS_21, HLHS_22, HLHS_35, HLHS_132 (siehe 2.1.1.1, Tabelle 1 Analyse „E“) – werden in diesem Abschnitt dargestellt. Mit Hilfe dieser Methode wurde in dieser Arbeit das betroffene Gewebe, das Herzgewebe, hinsichtlich krankheitsrelevanter Genvarianten analysiert.

3.2.1.1.1 Qualitätskontrolle und Datenfilterung

Eine primäre Qualitätskontrolle, Alignment, und Varianten-*Calling* der Rohdaten der WES-Analyse (Details siehe 2.2.1.7) erfolgte mittels Varbank-Pipeline v.2.3 (<https://varbank.ccg.uni-koeln.de/>) durch Bioinformatiker des CCG. Dies resultierte in einer mittleren *Reads Coverage* der gemappten Sequenzen der neun analysierten HLHS-Fälle von 73 *Reads*. Anschließend erfolgte in dieser Arbeit mithilfe der Filter-Oberfläche der Varbank-Pipeline v.2.3 eine weitere Qualitätskontrolle und Filterung der WES-Daten nach der in Abschnitt 2.2.1.7 dargestellten Filter-Pipeline (Abbildung 8). Die resultierende Anzahl an Varianten bzw. assoziierten Genen von jedem Filter-Vorgang ist als Übersicht in Abbildung 19 dargestellt.

Das Filtern nach Parametern des **QC-Filters** (Details siehe 2.2.1.7) resultierte in 3156 Varianten in 2645 Genen in allen neun HLHS-Patienten zusammen betrachtet (Abbildung 19). Um stringent hinsichtlich seltener Genvarianten zu filtern, wurde mit einer $AF \leq 0,01\%$ als Filterkriterium des **AF-Filters** gefiltert (siehe 2.2.1.7), was in 908 Varianten in 825 Genen resultierte (Abbildung 19). Die anschließende Filterung nach **Mutations-Typ**, um proteinstrukturverändernde Genvarianten und somit potenziell krankheitsrelevante Mutationen zu selektieren, ergab 735 seltene funktionelle Varianten in 685 Genen (Abbildung 19).

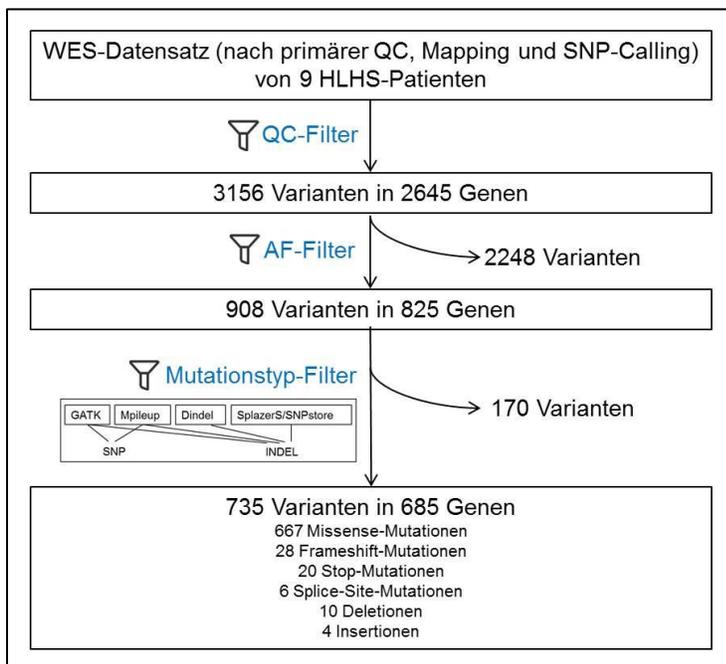


Abbildung 19 Prozessierung der WES-Daten von IAS-Gewebe von neun Patienten mit HLHS mit resultierender Anzahl an Varianten bzw. assoziierten Genen von jedem Filter-Vorgang. QC-Filter: Qualitätskontroll-Filter; AF-Filter: Filter nach Allelfrequenz (0,01 %); Mutationstyp-Filter: Filter nach Varianten, die die Proteinstruktur beeinflussen (Details siehe 2.2.1.7).

3.2.1.1.2 Datenanalyse

Die nach Qualitätskontrolle und Datenfilterung resultierenden 735 funktionellen Genvarianten in 685 Genen (als HLHS-WES-Datensatz im Folgenden bezeichnet) bestanden, entsprechend der Varbank-Pipeline v.2.3-Annotation der Mutationskonsequenz, aus 667 Missense-Mutationen, 28 Frameshift-Mutationen, 20 Stop-Mutationen, zwei Start-Loss-Mutationen, sechs Splice-Site-Mutationen, als auch 10 Deletionen und vier Insertionen, die keine Verschiebung des *open reading frame* (ORF) bewirkten (Abbildung 19). In Abbildung 20 sind die mittels WES detektierten Genvarianten pro Patient als Übersicht dargestellt. Im Durchschnitt konnte pro Patient eine Anzahl von 91,2 seltenen bzw. bisher unbekanntem Genvarianten detektiert werden. Bei allen Patienten konnten mindestens zwei Frameshift-Mutationen, die mit einem vorzeitigen Stopkodon einhergehen, detektiert werden. Mit Ausnahme von HLHS_4 traten bei allen Patienten mindestens zwei Stop-Mutationen auf. Start-Loss-Mutationen, d.h. Initiatorcodon-Mutationen, wurden bei den Patienten HLHS_21 und HLHS_22 (monozygote Zwillinge) detektiert. Splice-Site-Mutationen, sowie *inframe*-Deletionen (ohne Auswirkung auf den ORF) konnten bei fünf von neun Patienten detektiert werden. Drei Patienten zeigten *inframe*-Insertionen. Alle detektierten Mutationen waren heterozygot ausgeprägt, mit Ausnahme von einer hemizygoten Missense-Mutation in *ACE2* bei HLHS_4 (siehe Abschnitt 3.3.2). Des Weiteren konnte eine *compound* heterozygote Mutation in *DOCK1* beim Patienten HLHS_132 detektiert werden. Die Ergebnisse hierzu werden in Abschnitt 3.3.1 aufgeführt.

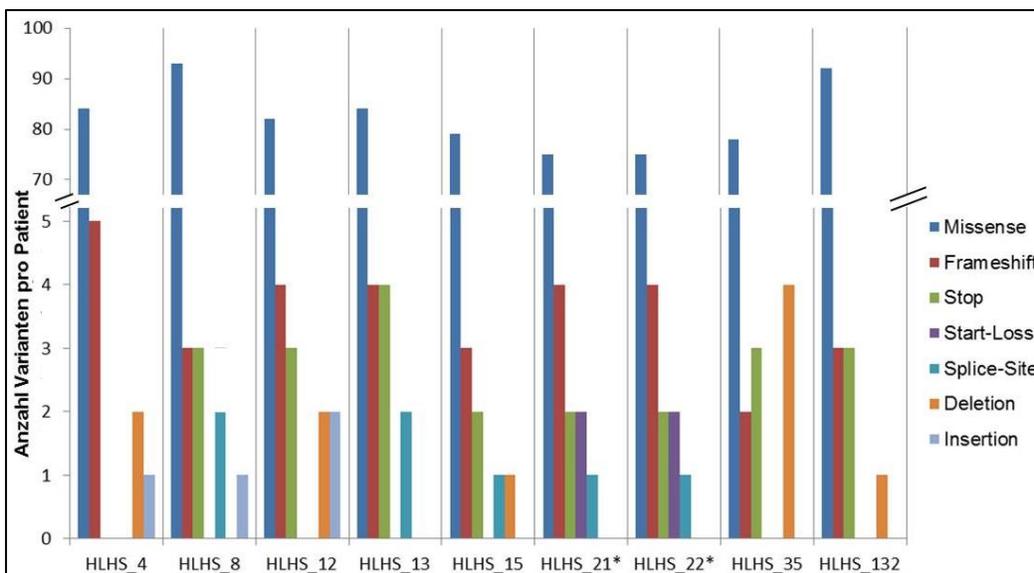


Abbildung 20 Übersicht der mittels WES-Analyse im IAS-Gewebe detektierten Genvarianten pro HLHS-Patient. Auf der horizontalen Achse sind die Probenbezeichnungen der Patienten aufgeführt. Die Proben der monozygoten Zwillinge sind mit einem Sternchen markiert. Die vertikale Achse stellt die Anzahl der Genvarianten pro Patient dar. Die parallelen Linien in der vertikalen Achse markieren die Unterbrechung des Achsenintervalls. Farbenzuordnung der Mutationen (Mutationskonsequenz nach Varbank-Pipeline v.2.3): blau = Missense, rot = Frameshift, grün = Stop, violett = Start-Loss, türkis = Splice-Site, orange = Deletion, blass-blau = Insertion.

Im IAS-Gewebe von neun Patienten mit HLHS konnten insgesamt 56 seltene bzw. bisher unbekannte LOF-Mutationen³⁰, d.h. Mutationen, die zu einem Genprodukt mit verringerter Funktion oder Funktionsverlust führen, detektiert werden (Tabelle 23). Alle LOF-Mutationen konnten nach visueller *Reads*-Analyse (siehe 2.2.1.7) als valide eingestuft werden. Der Abgleich der zuvor nach Varbank-Pipeline v.2.3 annotierten LOF-Mutationen wurde mithilfe der VEP-Datenbank (Variant Effect Predictor, Ensembl (1000G-rel77)) hinsichtlich der Mutationstyp-Annotationen abgeglichen. Die Mutationstyp-Annotationen von Stop-, Start-Loss- und Frameshift-Varianten stimmte mit denen der VEP-Datenbank überein und wurden daher als glaubwürdig eingestuft (Tabelle 23). Die nach Varbank-Pipeline v.2.3 annotierten sechs Splice-Site-Mutationen wurden nach VEP-Annotation in folgende Unterkategorien eingeteilt: drei Splice-Donor-Varianten (5'-Ende des Introns), eine Splice-Akzeptor-Variante (3'-Ende des Introns) und zwei Splice-Region-Varianten, die sich an Position -1 im Exon befanden (Tabelle 23).

Bei den identifizierten LOF-Mutationen im IAS-Gewebe von HLHS-Patienten handelte es sich hauptsächlich um bisher unbekannte bzw. durch ExAC nicht-annotierte Mutationen (Tabelle 23). Lediglich sechs der 56 LOF-Mutationen ließen sich zuvor schon nach ExAC-Datenbank identifizieren (jeweils AF < 0,01%). Nach prädiktiver Mutationskonsequenz durch Polyphen-2 bzw. Mutation Taster [268] (siehe 2.1.3.2) konnten sieben Mutationen als *probably damaging* eingestuft werden (Tabelle 23). Dazu zählten folgende Mutationen: Stop-Mutationen (*IFNA7*, *KCND1*, *LCLAT1*, *LPXN*, *TNNC2*), eine Splice-Region-Mutation in *TDRD3* und eine Splice-Donor-Mutation in *VPS13B*.

³⁰ umfasst Frameshift-, Stop-, Start-Loss- und Splice-Site-Mutationen

Tabelle 23 LOF-Mutationen, die im IAS-Gewebe von neun HLHS-Patienten detektiert wurden

Typ	ID	Gen (HGNC)	NM-Nummer	Position CHR (GRCh37/hg19)	Mutation (Konsequenz Proteinebene)	VEP	ExAC AF	PPH/M	verif.
Stop-Mutation	132	<i>ATP10B</i>	NM_025153.2	chr5:160097475	c.670C>T (p.Q224*)	Stop	-	-	
	8	<i>C5</i>	NM_001735.2	chr9:123753544	c.2866C>T (p.R956*)	Stop	-	-	
	12	<i>CCDC125</i>	NM_176816.3	chr5:68616124	c.244C>T (p.Q82*)	Stop	-	-	ja
	21/22	<i>DNAH8</i>	NM_001206927.1	chr6:38854585	c.8242C>T (p.R2748*)	Stop	-	-	ja
	12	<i>GALE</i>	NM_001127621.1	chr1:24122747	c.882C>A (p.Y294*)	Stop	-	-	
	13	<i>GDPD5</i>	NM_030792.6	chr11:75173864	c.188G>A (p.W63*)	Stop	-	-	ja
	132	<i>IFNA7</i>	NM_021057.2	chr9:21202105	c.60C>A (p.C20*)	Stop	-	D (M)	
	35	<i>IMPAD1</i>	NM_017813.4	chr8:57905838	c.307G>T (p.E103*)	Stop	-	-	ja
	15	<i>KCND1</i>	NM_004979.4	chrX:48826264	c.415C>T (p.R139*)	Stop	-	D (M)	
	8	<i>LCLAT1</i>	NM_182551.3	chr2:30863141	c.901C>T (p.Q301*)	Stop	-	D (M)	ja
	13	<i>LCNL1</i>	NM_207510.3	chr9:139879261	c.293G>A (p.W98*)	Stop	$2,2 \times 10^{-5}$	-	
	35	<i>LPXN</i>	NM_004811.2	chr11:58295014	c.1074C>A (p.C358*)	Stop	8×10^{-6}	D (M)	ja
	13	<i>LRRIQ1</i>	NM_001079910.1	chr12:85450720	c.2149G>T (p.E717*)	Stop	-	-	
	15	<i>MADD</i>	NM_003682.3	chr11:47297586	c.796C>T (p.R266*)	Stop	-	P (M)	ja
	21/22	<i>RADIL</i>	NM_018059.4	chr7:4855009	c.2039G>A (p.W680*)	Stop	-	-	ja
	12	<i>SCYL3</i>	NM_181093.2	chr1:169824956	c.1455C>G (p.Y485*)	Stop	-	-	ja
	13	<i>SMO</i>	NM_005631.4	chr7:128850348	c.1611G>A (p.W537*)	Stop	-	-	ja
	8	<i>TNNC2</i>	NM_003279.2	chr20:44453152	c.181G>T (p.E61*)	Stop	-	D (M)	ja
35	<i>USP24</i>	NM_015306.2	chr1:55561046	c.6085C>T (p.R2029*)	Stop	-	-	ja	
132	<i>ZNF418</i>	NM_133460.1	chr19:58438366	c.1183G>T (p.E395*)	Stop	-	-	ja	
Start Loss	21/22	<i>ASPG</i>	NM_001080464.2	chr14:104552108	c.1A>G (p.M1V)	Start	$6,4 \times 10^{-5}$	-	ja
	21/22	<i>LYPD3</i>	NM_014400.2	chr19:43969722	c.2T>C (p.M1T)	Start	-	B	ja
Splice-Site-Mutation	8	<i>ADCY10</i>	NM_018417.4	chr1:167825402	c.2171+1G>C	Splice-D	-	-	
	15	<i>GMFG</i>	NM_004877.2	chr19:39819640	c.357G>A (p.Lys119Lys)	Splice-R/synonym	8×10^{-6}	-	
	13	<i>NCAN</i>	NM_004386.2	chr19:19335785	c.779-1G>T	Splice-A	-	-	
	8	<i>SNX4</i>	NM_003794.2	chr3:125216184	c.597+1G>A	Splice-D	1×10^{-6}	-	ja
	13	<i>TDRD3</i>	NM_001146070.1	chr13:61018882	c.192G>T (p.K64N)	Splice-R/Missense	-	D	
	21/22	<i>VPS13B</i>	NM_152564.4	chr8:100519432	c.4224+1G>C	Splice-D	8×10^{-6}	D (M)	
Frameshift-Mutation	8	<i>ABLIM2</i>	NM_001130083.1	chr4:7986574	c.1724delG(p.G576Afs*21)	FS	-	-	ja
	13	<i>C19orf44</i>	NM_032207.2	chr19:16612151-16612152	c.548_549delAC (p.S185Pfs*2)	FS	-	-	
	13	<i>C21orf59</i>	NM_021254.2	chr21:33975571-33975581	c.557_568insAGGCAGAGGC (p.Q190Gfs*26)	FS	-	-	
	13	<i>C5orf60</i>	NM_001142306.1	chr5:179069992	c.558_559insG (p.S188Vfs*194)	FS	-	-	
	12	<i>DNAH3</i>	NM_017539.1	chr16:21151869	c.683_684insA (p.S229Ifs*22)	FS	-	-	ja

Typ	ID	Gen (HGNC)	NM-Nummer	Position CHR (GRCh37/hg19)	Mutation (Konsequenz Proteinebene)	VEP	ExAC AF	PPH/M	verif.
Frameshift-Mutation	8	FAM65A	NM_001193524.1	chr16:67575847-67575853	c.1200_1206delCAGCCCA (p.P402Sfs*59)	FS	-	-	ja
	15	GFM2	NM_032380.3	chr5:74043549-74043550	c.575_576delGT (p.C192Ffs*13)	FS	-	-	
	35	GPD2	NM_001083112.2	chr2:157367420	c.387delG(p.D130Lfs*8)	FS	-	-	ja
	4	HIST1H2A G	NM_021064.4	chr6:27101213-27101214	c.363_364delITG (p.S123Pfs*19)	FS	-	-	ja
	13	IFT88	NM_175605.3	chr13:21212565	c.1508delT (p.F504Lfs*17)	FS	-	-	ja
	132	IRAK2	NM_001570.3	chr3:10206728-10206729	c.90_91delGT (p.E30Dfs*68)	FS	-	-	
	21/22	KLHL8	NM_020803.3	chr4:88084724-88084727	c.1807_1810delACTA (p.T603Afs*6)	FS	-	-	ja
	4	LRIT2	NM_001017924.2	chr10:85984557	c.424delC (p.L142Sfs*12)	FS	-	-	ja
	35	MORN3	NM_173855.4	chr12:122092232-122092233	c.410_411delAC (p.Y137*fs)	FS	-	-	
	15	NOTCH1	NM_017617.3	chr9:139417629-139417636	c.408_414delATCGTGCinsGGGGCTGAC (p.S137Gfs*140)	FS	-	-	ja
	4	NUP37	NM_024057.2	chr12:102494851	c.313delA (p.R105Dfs*16)	FS	-	-	ja
	21/22	OR1S2	NM_001004459.1	chr11:57971093-57971094	c.558_559delTC(p.L187Pfs*6)	FS	-	-	ja
	21/22	PEPD	NM_001166057.1	chr19:33882302	c.859delG (p.V287Wfs*107)	FS	-	-	
	12	PNMA1	NM_006029.4	chr14:74180269	c.71delG (p.G25Afs*4)	FS	-	-	ja
	12	PSKH2	NM_033126.1	chr8:87060905-87060908	c.941_945insCCCT (p.D316Pfs*44)	FS	-	-	
	4	RG9MTD3	NM_144964.2	chr9:37776372-37776373	c.814_815delTA (p.Y272Sfs*12)	FS	-	-	
	4	RNF207	NM_207396.2	chr1:6279306	c.1744delG (p.G582Efs*29)	FS	-	-	ja
	132	RTN4	NM_207521.1	chr2:55200971-55200972	c.2843_2844insT (p.Y949Lfs*2)	FS	-	-	
	15	SEC24D	NM_014822.2	chr4:119660413-119660416	c.2265_2268delTGGT (p.G756Kfs*11)	FS	-	-	
	12	STEAP2	NM_001040666.1	chr7:89854672	c.276delT (p.I93Yfs*12)	FS	-	-	ja
21/22	TMCO2	NM_001008740.3	chr1:40717197	c.480delT (p.E161Sfs*25)	FS	-	-		
8	VNN1	NM_004666.2	chr6:133013485	c.1065delA (p.V356Ffs*10)	FS	-	-		
132	ZNF561	NM_152289.2	chr19:9721985-9721986	c.349_350delCT (p.C118*fs)	FS	-	-		
132	SOSTDC1 ³¹	NM_015464.2	chr7:16502202	c.592delA (p.S198Afs*8)	FS	0,0034		ja	

ID: HLHS-Probennummer; Position CHR: chromosomale Position der Mutation; VEP: *Variant Effect Predictor* (VEP) Konsequenz (1000G-rel77); Stop: Stop Variante, Start: Initiatorcodon Variante; FS: Frameshift Variante; Splice-D: Splice-Donor Variante; Splice-A: Splice-Akzeptor Variante; Splice-R: Splice-Region Variante; ExAC AF: Allelfrequenz nach ExAC (Version 0.3, Feb. 2015); PPH: Polyphen-2; M: Mutation Taster, D: *probably damaging*; P: *possibly damaging*; D (M): *probably damaging* nach Mutation Taster; verif.: verifiziert durch Sanger-Sequenzierung; grau schattiert: zusätzlich zu den 56 LOF-Mutation mit AF ≤ 0,01 % wurde SOSTDC1 (AF = 0,34 %) in die Liste mit einbezogen, da dieses Gen nachfolgend in einer integrativen Analyse zu Genom und Transkriptom (siehe 3.3.2) analysiert wurde.

³¹ Die FS-Mutation in SOSTDC1 mit einer AF > 0,01 % (nicht dem AF-Filter aus Abschnitt 2.2.1.7 entsprechend) wurde in diese Liste mit eingefügt, da SOSTDC1 in einer integrativen Analyse zu Genom und Transkriptom (siehe 3.3.2) nach Filterung der WES-Daten mit dem Kriterium AF ≤ 5 % in die Analysen einbezogen wurde.

Analyse der Überlappung zu bereits in der Literatur bekannten Genen mit kardiovaskulärem Phänotyp

Der HLHS-WES-Datensatz zu den neun HLHS-Patienten sollte in dieser Analyse hinsichtlich einer Überlappung von bekannten Genen, die in Assoziation mit Herzfehlbildungen identifiziert wurden, untersucht werden. Dazu wurden kardiale Genlisten (als Herz-Genliste bezeichnet), im Vorfeld erstellt durch Dr. med. Marc-Phillip Hitz [283], verwendet (Details siehe 2.2.1.7). Die Überlappung dieser Herz-Genliste mit den 685 Genen des HLHS-WES-Datensatzes ist in Tabelle 24 dargestellt.

Tabelle 24 Übersicht der Überlappung der Herz-Genliste mit 685 Genen mit funktionellen Varianten im IAS-Gewebe von neun HLHS-Patienten

Überlappung „kardialer“ Gene mit 685 Genen mit funktionellen Varianten im IAS-Gewebe von neun HLHS-Patienten	
Gene „kardialer Phänotyp Maus“ 17 Gene von 685 Genen	Gene „kardialer Phänotyp Human“ 10 Gene von 685 Genen
<u>ADCY10</u> , APP, CNTNAP1, DICER1, DOK7, ECE2, FBN2, HAND1, <u>IFT88</u> ¹ , NOS1, PAPP2, PLEC, PTCH1, RYR1, RYR3, SLC6A9, SMTN	NAA15, ROR2 ¹ , ARSB, CREBBP ¹ , MLL, <u>NOTCH1</u> ¹ , NSD1, OFD1, SURF1, TTN syndrom-assoziiert, nicht-syndromal, CM

unterstrichen = LOF-Mutation, Index 1 = Mutation in diesem Gen verifiziert durch Sanger-Sequenzierung; rot = Gene, die bei Fällen nicht-syndromaler CHDs identifiziert wurden (Literaturrecherche PubMed-Datenbank); grün = Gene, die bei Fällen syndrom-assoziiierter CHDs identifiziert wurden; blau = Gene, die bei Fällen mit Kardiomyopathie (CM) identifiziert wurden.

Insgesamt **17 der 685 Gene** des HLHS-WES-Datensatzes zeigten eine Überlappung mit Genen mit **kardiovaskulärem Phänotyp im Mausmodell** nach MGI-Datenbank (Tabelle 24). Zwei dieser 17 Gene wiesen eine LOF-Mutation auf. Hierzu zählten (1) *ADCY10* mit einer heterozygoten Splice-Donor-Mutation bei HLHS_8, wobei Knockout-Mäuse dieses Gens eine erhöhte Herzfrequenz zeigten (MGI Ref ID: 103485), und (2) *IFT88* mit einer heterozygoten Frameshift-Mutation bei HLHS_13, welche mittels Sanger-Sequenzierung verifiziert werden konnte (Tabelle 23). Knockout-Mäuse wiesen hierbei eine Letalität in der frühen Gestationsphase auf [308]. Mit einer Missense-Mutation in *HAND1* (c.122A>G) im IAS-Gewebe eines HLHS-Patienten war zudem ein TF der Herzentwicklung [309, 310] durch eine funktionelle Genvariante betroffen.

Zehn der 685 Gene des HLHS-WES-Datensatzes waren mit einem **kardiovaskulären humanen Phänotyp** in der Literatur (NCBI PubMed-Datenbank) beschrieben. Da es sich bei den analysierten HLHS-Patienten in dieser Arbeit um nicht-syndromale CHDs handelte, wird im Folgenden lediglich auf die Assoziationen in der Literatur bekannter nicht-syndromaler Fälle genauer eingegangen.

Mutationen in *NAA15* und *NOTCH1* konnten zuvor schon bei Fällen **nicht-syndromaler CHDs** identifiziert werden. *NAA15* zeigte bei HLHS_12 und HLHS_21/22 heterozygote Missense-Mutationen (Abbildung 21). Bei HLHS_21/22 konnte eine bisher unbekannte Genvariante c.945G>C (p.Arg315Ser) und bei HLHS_12 eine seltene (AF = 0,00024) Genvariante c.2350A>G (p.Ile784Val) nachgewiesen werden. In der Studie von Zaidi et al. (2013) [74] ließen sich in diesem Gen bei zwei Patienten mit CHDs *de novo* LOF-Mutationen identifizieren. In *NOTCH1* konnte bei HLHS_15 eine bisher unbekannte komplexe INDEL-Mutation c.408_414delATCGTGCinsGGGGCTGAC (p.Ser137Glyfs*140), d.h. eine Frameshift-Mutation mit vorzeitigem Stopkodon, detektiert werden. Diese konnte mittels Sanger-Sequenzierung verifiziert werden (Tabelle 23). Auch in vorherigen Studien konnten *NOTCH1*-Mutationen in Zusammenhang mit LS-CHDs, inklusive dem HLHS, nachgewiesen werden [196, 225, 226].

Mutationen in sechs Genen – *ROR2*, *ARSB*, *CREBBP*, *MLL*, *NSD1* und *OFD1* – wurden zuvor bei Fällen mit **syndrom-assoziierten CHDs** identifiziert. In *SURF1* und *TTN* konnten zuvor schon Mutationen bei Fällen mit **Kardiomyopathie** nachgewiesen werden.

Analyse der Mutationstoleranz identifizierter Gene

Zur weiteren Interpretation und Analyse der WES-Daten wurde das **RVIS-Scoring-System** (RVIS: *Residual Variation Intolerance Score*) nach Petrovski et al. (2013) [264] (Details siehe 2.2.1.7) verwendet. Mehr als ein Drittel der 685 Gene mit funktionellen Genvarianten, die im IAS-Gewebe der neun HLHS-Patienten detektiert werden konnten, zeigten mit einem RPS (RVIS-Perzentil-Score) ≤ 25 eine geringe Mutationstoleranz. Letztgenannte Gene wiesen somit nach RVIS-Scoring-System eine hohe Suszeptibilität gegenüber funktionellen Varianten auf. Innerhalb dieser mutationsintoleranten Genen mit RPS ≤ 25 ließen sich neun Gene identifizieren, die in 2–3 der 9 analysierten HLHS-Patienten funktionelle Genvarianten aufwiesen. In diesen Genen mit geringer Mutationstoleranz konnten somit in mehr als einem HLHS-Patienten Mutationen detektiert werden und eine Übersicht der identifizierten Genvarianten ist in Abbildung 21 aufgeführt.

Gen (HGNC)	RPS	HLHS_4		HLHS_8		HLHS_12		HLHS_13		HLHS_15		HLHS_21*		HLHS_22*		HLHS_35		HLHS_132	
		Typ	PPH	Typ	PPH	Typ	PPH	Typ	PPH	Typ	PPH	Typ	PPH	Typ	PPH	Typ	PPH	Typ	PPH
ANKRD11	0,62		P														B		
EP400	1,01				-		-												
NAA15	9,72						B					D		D					
NCAN	6,15											D		D					
PTCH1	12,43						-												
RYR1	0,03				-												-		D
SMURF2	15,8				D														D
USP24	1,23																	-	P
ZNF318	0,69						D				D								

Farblegende:

	Missense
	Stop
	Splice-Akzeptor
	Gen mit kardiovaskulärem Phänotyp im Mausmodell

PPH-Legende:

B	<i>benign</i>
P	<i>possibly damaging</i>
D	<i>probably damaging</i>
-	kein PPH-Wert annotiert

Abbildung 21 Übersicht der Gene mit einem RPS ≤ 25 (RVIS-Perzentil-Score nach Petrovski et al. (2013) [264]) mit funktionellen Genvarianten, die im IAS-Gewebe von mindestens zwei Patienten mit HLHS detektiert werden konnten. Links sind die Gennamen (nach HGNC) und der jeweilige RPS-Wert (Werte zwischen 0 und 100, mit Werten nahe 0 = Gen mit geringer Mutationstoleranz) aufgeführt. Gelb hinterlegte Gene stellen Gene mit kardiovaskulärem Phänotyp im Mausmodell nach MGI-Datenbank dar. Pro Patienten (obere Zeile der Tabelle) sind jeweils der Mutationstyp (Typ) als Farbcode (siehe Farblegende) und der annotierte Polyphen-2-Wert (siehe PPH-Legende) pro Gen aufgeführt. Die Zwillingssproben (*) wiesen jeweils dieselben Mutationen auf und wurden daher in dieser Analyse als ein „zusammengefasster“ Patient betrachtet.

Drei der Gene mit Mutationen in zwei oder mehr HLHS-Patienten aus Abbildung 21 – *NAA15*, *PTCH1* und *RYR1* – zeigten einen kardiovaskulären Phänotyp nach MGI-Datenbank im Mausmodell. *NAA15*, war zusätzlich auch mit einem humanen kardialen Phänotyp assoziiert ([74] s.o.). Als *probably damaging* sind Mutationen in fünf der neun Gene nach Polyphen-2 kategorisiert (Abbildung 21). In *NCAN* und *USP24* konnte jeweils, zusätzlich zu einer Missense-Mutation, eine LOF-Mutation detektiert werden (Abbildung 21 und Tabelle 23 (Stop-Mutation in *USP24* wurde mittels Sanger-Sequenzierung verifiziert)).

Eine Analyse der RPS-Verteilung der 685 Gene des HLHS-WES-Datensatzes im Kontext mit RPS-Werten aller humanen Gene und bekannter Gene mit kardiovaskulärem Phänotyp (Herz-Genliste) ist in Abbildung 22 dargestellt.

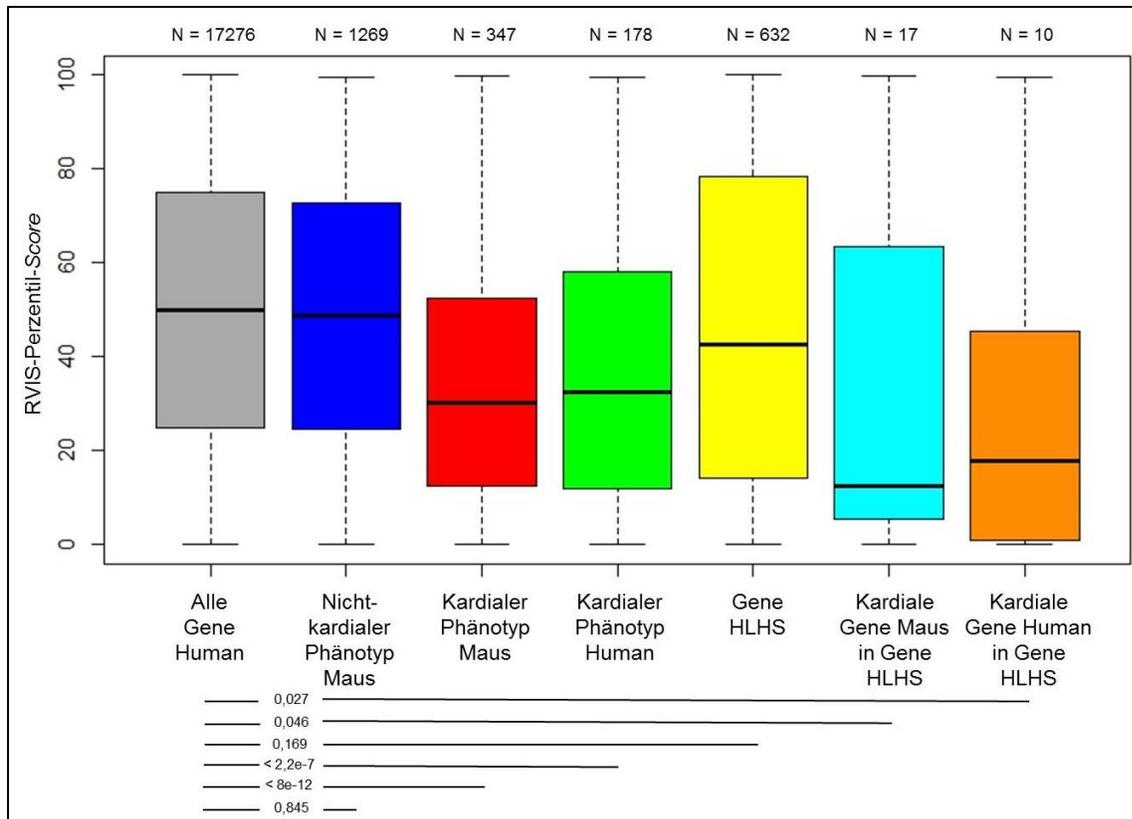


Abbildung 22 Verteilung des RVIS-Perzentil-Scores (RPS; nach Petrovski et al. (2013) [264]) der 685 Gene mit funktionellen Varianten des HLHS-WES-Datensatzes im Kontext mit RPS-Werten aller humanen Gene und bekannter Gene mit kardiovaskulärem Phänotyp (aus Herz-Genliste). Die RPS-Werte sind auf der vertikalen Achse dargestellt. Die Boxplots zu den einzelnen Gengruppen sind auf der horizontalen Achse dargestellt. Oberhalb der Boxplots ist die Anzahl (N) der Gene mit annotierten RVIS-Scores pro Gengruppe aufgeführt. Farbdarstellung Boxplots: grau = alle humanen Gene, blau = Gene ohne kardiovaskulärem Phänotyp im Mausmodell (Daten aus MGI-Datenbank), rot = Gene mit kardiovaskulärem Phänotyp im Mausmodell (Daten aus MGI-Datenbank), grün = Gene mit kardiovaskulärem humanen Phänotyp (Daten aus NCBI PubMed-Datenbank), gelb = Gene mit funktionellen Varianten im HLHS-WES-Datensatz, türkis = Gene mit kardiovaskulärem Maus-Phänotyp aus dem HLHS-WES-Datensatz, orange = Gene mit kardiovaskulärem humanen Phänotyp aus dem HLHS-WES-Datensatz. Unterhalb der Abbildung ist der p-Wert des Wilcoxon *Signed-Rank* Tests der jeweiligen Vergleiche einzelner Gengruppen dargestellt.

Verglichen mit RPS-Werten aller humaner Gene zeigten Gene ohne kardiovaskulärem Phänotyp im Mausmodell keinen signifikanten Unterschied ($p = 0,845$) im Wilcoxon *Signed-Rank* Test (Abbildung 22). Dagegen wiesen Gene der Herz-Genliste, d.h. Gene mit kardiovaskulärem Phänotyp im Mausmodell ($p < 8 \times 10^{-12}$) und Gene mit kardiovaskulärem humanen Phänotyp ($p < 2,2 \times 10^{-7}$), signifikant geringere RPS-Werte als die Gesamtheit aller humaner Gene. Mit einem mittleren RPS-Median von 31,1 zeigten letztere kardiale Gengruppen eine geringe Mutationstoleranz auf (Abbildung 22). Gene mit funktionellen Varianten bei HLHS-Patienten (HLHS-WES-Datensatz) zeigten keine signifikante Anreicherung ($p = 0,169$) mutationsintoleranter Gene. Betrachtete man jedoch Gene mit einem kardialen Phänotyp innerhalb des HLHS-WES-Datensatzes (Tabelle 24) die 17 Gene mit kardiovaskulärem Phänotyp im Mausmodell ($p = 0,046$) und die 10 Gene mit kardiovaskulärem humanen Phänotyp

($p = 0,027$), so zeigten diese eine signifikante Anreicherung von Genen mit geringer Toleranz gegenüber funktionellen Varianten ((RPS-Median < 20) Abbildung 22).

Eine **Anreicherungsanalyse** mittels WebGestalt-Tool ergab beim HLHS-WES-Datensatz keine signifikante Anreicherung bei einem Signifikanzlevel von $p \leq 0,001$ in der *Gene Ontology* (GO) Analyse.

Eine **Netzwerkanalyse** mittels STRING-Software (Details siehe 2.2.5.3), die Protein-Protein-Interaktions-Netzwerke darstellt, wurde anhand der 685 Gene durchgeführt. In Abbildung 23 sind ausschließlich die Netzwerke dargestellt, die aus mehr als fünf Interaktionspartnern bestanden. Hierbei lagen folgende Interaktionsnetzwerke vor: (1) Interaktionsnetzwerk aus Entwicklungsprozess-assoziierten Genprodukten, die in Notch- (*NOTCH1*) und Wnt- (*CTNNB1*) Signalwegen sowie in der transkriptionellen Regulation (wie *NCOR2*, *CREBBP*, *MED13L*, *MED1*, *NR5A2*) involviert sind, und (2) ein Interaktionsnetzwerk von Proteinen, die in der post-transkriptionellen Regulation eine Rolle spielen, wie der Prä-mRNA-Prozessierung (*SNRPB*, *SRRM1*, *SRRM2*, *PRPF6*) sowie dem nukleozytoplasmatischen RNA-Transport (*HNRNPUL1*, *NXF1*, *NUP37*, *NUP88*).

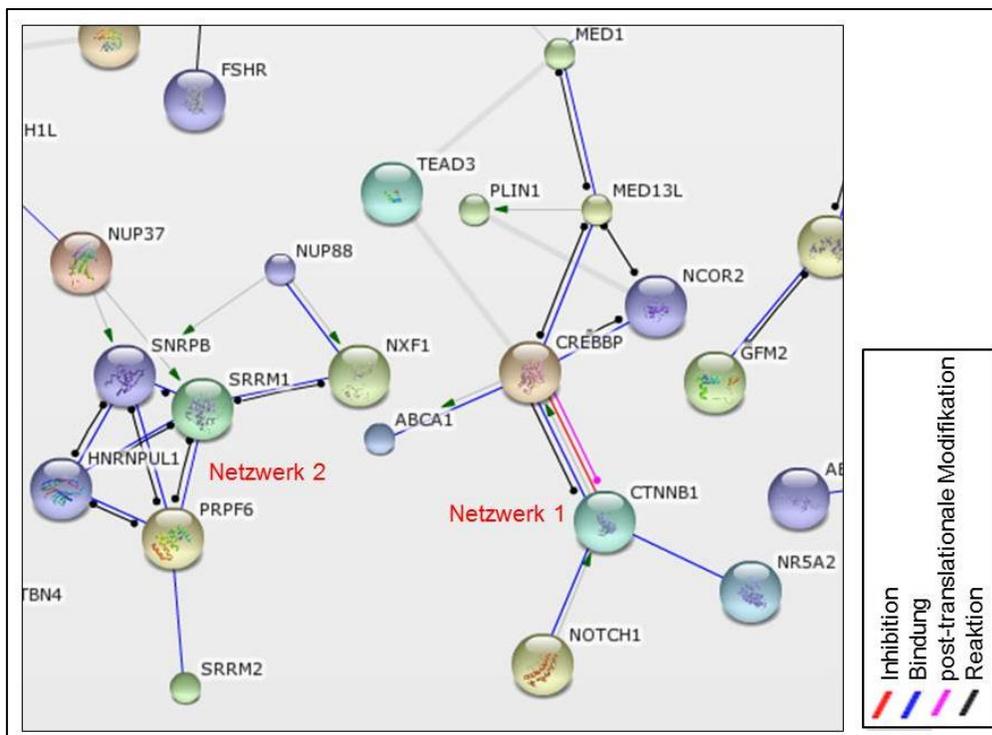


Abbildung 23 Netzwerkanalyse der Protein-Protein-Interaktionen nach STRING-Software von 685 Genen mit funktionellen Genvarianten im IAS-Gewebe von neun HLHS-Patienten durchgeführt. Es sind ausschließlich Netzwerke dargestellt, die aus mehr als fünf Interaktionspartnern bestehen (Netzwerk 1 und 2). Farbcode Protein-Interaktionen: rot = Inhibition, blau = Protein-Bindung, rosa = post-translationale Modifikation, schwarz = gemeinsame Beteiligung der Proteine in Reaktionen.

3.2.1.1.3 Verifizierung

Zur Verifizierung wurden 30 der 56 LOF-Mutationen und 22 der 667 Missense-Mutation mittels Sanger-Sequenzierung (siehe 2.2.1.6, Primer siehe Abschnitt 8 Tabelle 31) resequenziert. In Tabelle 25 ist eine Übersicht der Anzahl verifizierter Mutationen pro Mutationstyp dargestellt. Die jeweils verifizierten Mutationen sind auch in Tabelle 23 aufgeführt. Es konnten alle resequenzierten Mutationen (insgesamt 52 Mutationen) verifiziert werden. Der nach Angaben aus Abschnitt 2.2.1.7 gefilterte HLHS-WES-Datensatz konnte somit als valide eingestuft werden. Bei den Fällen HLHS_4, HLHS_21, HLHS_22 und HLHS_132 lag zusätzlich zur DNA aus IAS-Gewebe auch DNA aus pB-Proben der Indexpatienten sowie deren Eltern vor. Im Zuge der Resequenzierung wurden diese Proben verwendet, um mögliche *de novo* bzw. somatische (IAS-spezifische) Mutationen identifizieren zu können. Jede der bisher analysierten Varianten war jedoch auch im Keimbahnmaterial des Indexpatienten detektierbar und somit nicht somatisch. Die Varianten lagen jeweils heterozygot bei Mutter oder Vater vor, sodass keine *de novo* Mutationen bei den bisher analysierten Eltern-Indexpatient-Trios vorlagen.

Tabelle 25 Übersicht der Anzahl der mittels Sanger-Sequenzierung verifizierter Mutationen aus der WES von IAS-DNA von neun Patienten mit HLHS

Mutationstyp	Anzahl Mutationen	Anzahl verifiziert/Anzahl resequenziert
Stop	20	12/12
Start-Loss	2	2/2
Splice-Site	7	1/1
Frameshift	29	15/15
Missense	667	22/22

Mutationstyp: nach Annotation durch Varbank-Pipeline v.2.3; Anzahl Mutationen: Anzahl Mutationen des jeweiligen Mutationstyp innerhalb der gefilterten 735 Varianten (inklusive *SOSTDC1*: 736 Varianten); Anzahl verifiziert/Anzahl resequenziert: Anzahl verifizierbarer Mutationen pro resequenzierter Mutation mittels Sanger-Sequenzierung

Zusammenfassend konnten mittels WES-Analyse im IAS-Gewebe von neun Patienten mit HLHS insgesamt 685 Gene mit potenziell krankheitsrelevanten Genvarianten detektiert werden. 50 der 56 identifizierten LOF-Mutationen stellten bisher unbekannte neue Genvarianten dar. Mehr als ein Drittel der 685 Gene zeigten eine geringe Mutationstoleranz ($RPS < 25$). 17/685 Gene waren mit einem kardiovaskulären Phänotyp im Mausmodell assoziiert und 10/685 Gene waren bereits zuvor bei Fällen mit kardiovaskulären Erkrankungen identifiziert worden.

3.2.1.2 Kombinierte Keimbahn- und somatische Analyse in einer Familie mit HLHS identifiziert *ANKRD11* und *BHLHA15* mit potenziellen herzwegebespezifischen Mutationen

Bisher liegen nur wenige Studien zu somatischen Mutationen bei Patienten mit kardialen Defekten vor, deren Kausalität bisher noch ungeklärt ist. Um daher mögliche somatische Mutationen zu analysieren, wurden WES-Daten von Keimbahn- und IAS-Gewebe des HLHS-Falls HLHS_4, sowie dessen Eltern (Familie_4, siehe 2.1.1 Tabelle 2 bzw. 2.2.1.7 Tabelle 13), hinsichtlich IAS-spezifischer Mutationen analysiert.

3.2.1.2.1 Qualitätskontrolle und Datenfilterung hinsichtlich IAS-spezifischer Mutationen

Die Filterung hinsichtlich IAS-spezifischer Genvarianten (Details siehe 2.2.1.7) der WES-Rohdaten der IAS-Probe des HLHS-Falls HLHS_4 und der pB-Proben des entsprechenden Eltern-Index-Trios³² (Familie_4, siehe 2.2.1.7 (Tabelle 13)) resultierte in 11 heterozygoten Genvarianten, die nach Varbank-Filteroberfläche ausschließlich im IAS-Gewebe des HLHS-Patienten detektiert werden konnten (siehe Abschnitt 8 Tabelle 38). Nach visueller Analyse der *Reads Coverage* (siehe 2.2.1.7) zeigten sieben Varianten zusätzlich zur Mutation im Herzgewebe, auch in der Keimbahn des Patienten und/oder eines Elternteils die Mutation, sodass diese ausgeschlossen wurden. Vier der 11 Varianten wiesen bei Analyse der Keimbahn-Proben eine zu geringe bzw. keine *Coverage*, sodass keine Aussage getroffen werden konnte (siehe Abschnitt 8 Tabelle 38). Zudem konnten generell Unterschiede in der *Reads Coverage* der Exom-Daten von Herzgewebe-Proben, die mittels SeqCap EZ Human Exome Library v2.0 Enrichment-Kit (Roche NimbleGen) und Keimbahn-Proben, die mittels Agilent SureSelect Human All Exon 50MB Kit (Agilent Technologies) prozessiert wurden, detektiert werden. Die analysierten chromosomalen Positionen der 11 Varianten zeigten tendenziell bei Daten, die auf dem „Agilent“- Enrichment-Kit basierten, eine geringere *Coverage* als bei Daten, die auf dem „NimbleGen“-Enrichment-Kit basierten.

Als putativ somatische Mutationen, die ausschließlich das Herzgewebe betrafen, konnten final je eine Missense-Mutation c.6464C>A (p.P2155H) in *ANKRD11* (NM_013275.5) sowie eine

³² WES-Rohdaten stammen von einer anderen Sequenzierungsplattform mittels Agilent SureSelect Human All Exon 50MB Kit (siehe 2.2.1.7 Tabelle 13)

Missense-Mutation c.6G>T (p.K2N) in *BHLHA15* (NM_177455.3) identifiziert werden. Beide Mutationen wurden als *possibly damaging* durch Polyphen-2 eingestuft und nach ExAC-Datenbank als bisher unbekannt annotiert (Abbildung 24 C).

3.2.1.2.2 Verifizierung identifizierter somatischer Mutationen der WES-Analyse mithilfe der Sanger-Sequenzierung

Um diese potenziell somatischen Mutationen zu verifizieren, wurde eine Sanger-Sequenzierung (siehe 2.2.1.6) der IAS-Probe des HLHS-Falls HLHS_4 und der pB-Proben des entsprechenden Eltern-Index-Trios (Familie_4, siehe 2.2.1.7 (Tabelle 13)) durchgeführt (Abbildung 24). Die Sanger-Sequenzierungs-Analyse bestätigte, dass die Mutationen in *ANKRD11* und in *BHLHA15* nicht in der Keimbahn, d.h. in den pB-Proben von Indexpatient und Eltern, auftraten (siehe Abbildung 24). Die zuvor mittels WES-Methode im IAS-Gewebe des Patienten HLHS_4 detektierten Mutationen ließen sich jedoch nicht mithilfe der Sanger-Sequenzierungstechnologie resequenzieren. Auch unter Verwendung mehrerer unterschiedlicher Primerpaare (siehe Abschnitt 8 Tabelle 31) war eine Detektion dieser Mutationen mittels Sanger-Sequenzierung nicht möglich.

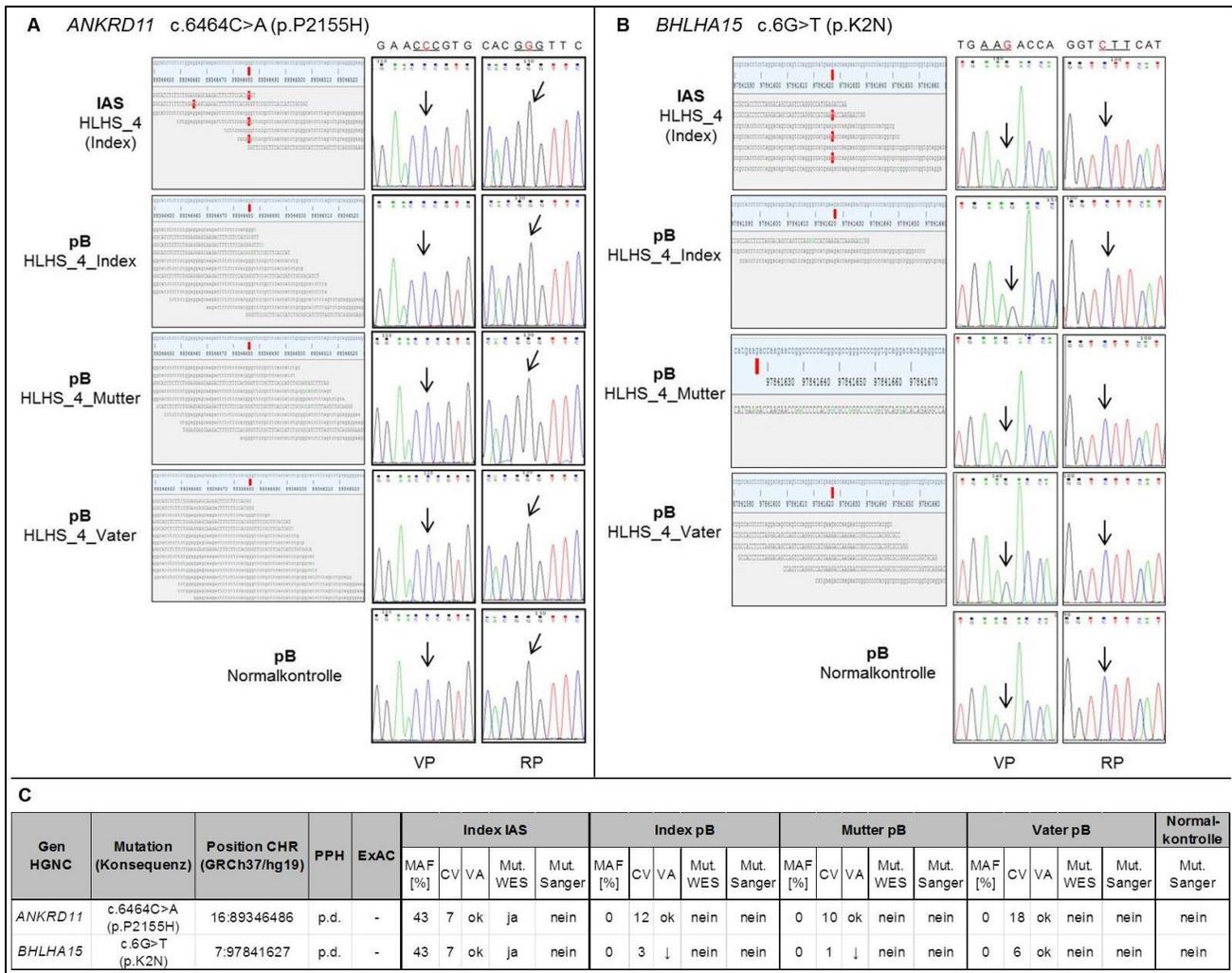


Abbildung 24 Untersuchung von somatischen (IAS-spezifischen) Mutationen bei dem Patienten HLHS 4 im familiären Kontext. In Abbildung A ist die Mutation in *ANKRD11* (c.6464C>A (p.P2155H)) und in Abbildung B die Mutation in *BHLHA15* (c.6G>T (p.K2N)) dargestellt. Jeweils links sind die Reads der WES-Analyse als Screenshot der Varbank-Pipeline v.2.3 der einzelnen Proben (links beschriftet) dargestellt, mit der jeweiligen Position der Mutation in rot oberhalb der Reads markiert. In Großbuchstaben sind Vorwärts-Sequenzen und in Kleinbuchstaben Rückwärtssequenzen dargestellt (nur schematisch zu erkennen). Da diese Screenshots der Reads Coverage in ihrer kleinen Darstellung lediglich einer schematischen Übersicht dienen, sind in der Tabelle (C) die Anzahl der Reads (CV) sowie die visuelle Reads-Analyse (VA) pro Probe aufgeführt. In den jeweiligen pB-Proben waren die Genvarianten jeweils in der visuellen Reads-Analyse (Screenshots) nicht detektierbar, sodass von IAS-Gewebespezifischen Varianten ausgegangen wurde. In A und B sind pro Probe die Elektropherogramme der Sanger-Sequenzierungen mit jeweils der Sequenz entsprechend des Vorwärts-Primers (VP) und Rückwärts-Primers (RP) dargestellt. Mit den Pfeilen sind die theoretischen Positionen der Mutationen darin markiert. Rote Buchstaben stellen die mutierte Base dar, unterstrichene Buchstaben die mutierte Kodonsequenz der Aminosäuren. Zusätzlich wurden die Mutationen in einer unabhängigen DNA-Normalkontrolle analysiert. C: Ergebnistabelle der WES-Analyse und Sanger-Sequenzierung. PPH: Polyphen-2, p.d. = *probably damaging*, ExAC: AF nach ExAC (Version 0.3, Feb. 2015). MAF: mutierte Allelfrequenz; CV: Reads Coverage (Anzahl Reads); VA: Visuelle Reads-Analyse; Mut. WES: Mutation in WES-Analyse vorhanden?; Mut. Sanger: Mutation in Sanger-Sequenzierung vorhanden?; ja = Mutation detektiert; nein = Mutation nicht detektiert; ok = mehr als 6 Reads; sowohl Vorwärts- und Rückwärts-Sequenzen vorhanden; ↓ = weniger als 6 Reads.

Die WES-Methode ließ somit zwar auf das mögliche Vorhandensein von somatischen, nur im Herzgewebe vorkommenden, Mutationen in *ANKRD11* und *BHLHA15* schließen, deren Detektion mittels Sanger-Sequenzierungstechnologie war bisher jedoch nicht möglich.

3.2.2 DNA-Methylierungsanalysen an Herzgewebe von Patienten mit verschiedenen kongenitalen Herzfehlbildungen

Die DNA-Methylierung konnte bisher mit verschiedenen Erkrankungen, wie Krebs [311, 312], Diabetes [313, 314], oder psychische Erkrankungen [315, 316] assoziiert werden. Zur DNA-Methylierung bei Patienten mit Herzerkrankungen liegen nur wenige Studien vor (siehe 1.4.2), wobei jedoch hauptsächlich adulte Herzgewebeproben analysiert wurden [177, 178]. In der vorliegenden Arbeit wurden daher die DNA-Methylierungsmuster von Kindern mit CHDs, insbesondere des HLHS, untersucht, um mögliche Aberrationen in der DNA-Methylierung zu identifizieren.

Herzgewebeproben von Patienten mit CHDs, speziell des HLHS, wurde in dieser Arbeit mittels Infinium HumanMethylation450 BeadChip Array (Illumina; 450K-Array abgekürzt) hinsichtlich des Methylierungsstatus von 485577 Loci analysiert. Hierbei wurden sechs Teilprojekte durchgeführt, in denen unterschiedliche Fragestellungen im Zuge der Analyse verschiedener Herzgewebe-Proben bearbeitet wurden. In Tabelle 26 sind die Teilprojekte, statistischen Analysen und pro Teilprojekt analysierte Herzgewebe-Proben kurz zusammengefasst.

Tabelle 26 Übersicht der DNA-Methylierungsanalysen mittels 450K-Array der verschiedenen Teilprojekte P1 – P5 mit Fragestellung und eingeschlossenen Herzgewebeproben (Diagnose Patientenproben und Herzgewebetyp) pro Teilprojekt

Teilprojekt-Nummer und Teilprojekt-Name	Fragestellung pro Teilprojekt	Herzgewebe-Proben	
		eingeschlossene Diagnosen	Typ
P1: HLHS-RF-Analyse	Liegt eine differenzielle DNA-Methylierung im Herzgewebe von HLHS-Patienten im Vergleich zu Rechtsherzfehlbildungen als Kontrollkohorte vor?	HLHS, TA, PAT, PAT & VSD	IAS
P2: Herz-Mapping-Analyse	Lassen sich verschiedene Herzregionen durch ein differenzielles DNA-Methylierungsprofil unterscheiden?	AVSD, DCM, HLHS, HLHS-TGA, LV-An, PAT, PAT & VSD, TA, normal	IAS, LA, RA, LV, RV, MV, TV, PA, Aorta, AoV, fetales HG
P3: Atrium-Ventrikel-Analyse	Unterscheiden sich Atrium- und Ventrikelgewebe in ihrer DNA-Methylierung und welche Gene sind differenziell methyliert?	AVSD, DCM, HLHS, HLHS-TGA, LV-An, PAT, PAT & VSD, TA	IAS, LA, RA, LV, RV
P4: Muskel-Gefäß-Analyse	Liegt eine differenzielle DNA-Methylierung in Herzmuskelgewebe im Vergleich zu Gefäßstrukturen vor? Welche Gen-Pathways zeigen eine differenzielle DNA-Methylierung?	DCM	IAS, LA, RA, LV, MV, TV, PA, Aorta, AoV
P5: HLHS-Fetal-Analyse	Welche Gene liegen im Herzgewebe von HLHS-Patienten im Vergleich zu gesundem fetalem Herzkontrollgewebe differenziell methyliert vor?	HLHS, normal	IAS, fetales HG

Detaillierte Angaben zu dem Herzgewebeproben-Typ und eingeschlossenen Diagnosen sind im Abschnitt 2.1.1.1 (Abbildung 5) zu finden; normal: keine Herzerkrankung; HG: Herzgewebe

Qualitätskontrolle

Insgesamt wurden 78 DNA-Proben prozessiert und ausgewertet. Für jedes der Teilprojekte erfolgte in der vorliegenden Arbeit eine Qualitätskontrolle, Normalisierung, Filterung der Daten und statistische Analyse nach Angaben aus Abschnitt 2.2.2.2.1. Keine der insgesamt 78 Hybridisierungen³³ zeigte eine LCR (*Loci Call Rate*) < 98 % (Tabelle 27), sodass nach diesem Ausschlusskriterium keine Proben exkludiert werden mussten. Der LCR-Median über alle Analysen gesehen betrug 99,96 %. Zur **Qualitätskontrolle der Bisulfit-Konversion** wurden die *Quality Control Bar Plots* der RnBeads-Analyse (Details siehe 2.2.2.2.1 und 2.2.5.1.1) überprüft. In Abbildung 25 ist ein *Quality Control Bar Plot* der BSK-Sonde „BS Conversion I U2“, einer unkonvertierten BSK-Sonde, dargestellt. Hierbei ist zu erkennen, dass die Probe HLHS_157 eine hohe Fluoreszenzsignalintensität (> 2000 FE) und somit eine Hybridisierung an die unkonvertierte BSK-Sonde aufwies. Somit lag eine unvollständige Bisulfit-Konversion dieser Probe vor. Die Probe HLHS_157 wurde daher aus der weiteren Analyse ausgeschlossen (Tabelle 27). Die restlichen Proben des in Abbildung 25 dargestellten Teilprojekts P1 (HLHS-RF-Analyse) zeigten Fluoreszenzsignalintensitäten im Bereich des erwarteten Hintergrundsignals (> 200 FE) und ließen daher auf eine effiziente Bisulfit-Konversion schließen. Alle weiteren Teilprojekte (P 2 – P 5) ergaben keinen Probenausschluss aufgrund einer ineffizienten Bisulfit-Konversion.

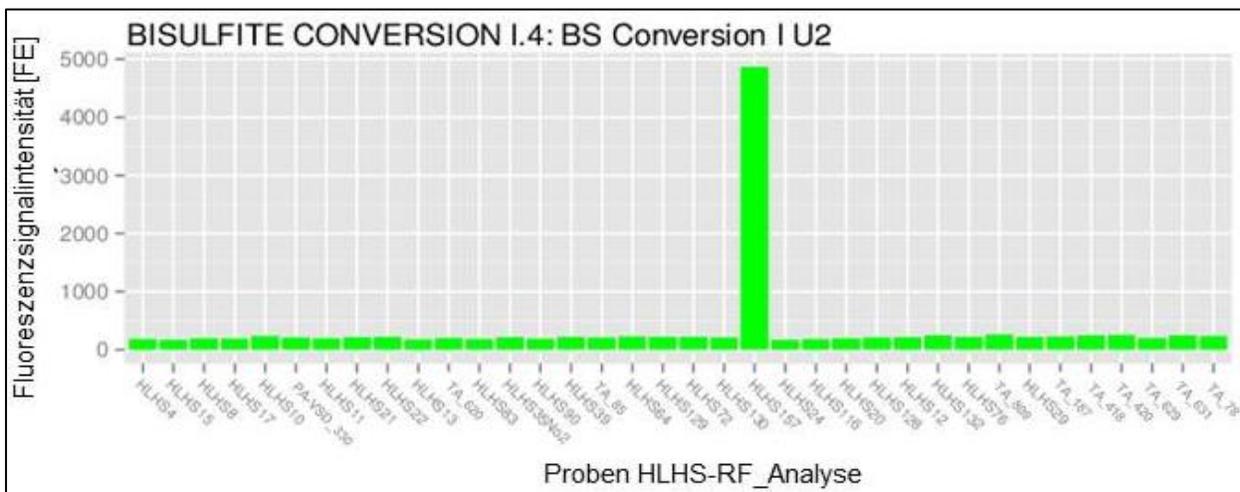


Abbildung 25 *Quality Control Bar Plot*-Darstellung der RnBeads-Analyse zur Qualitätskontrolle der Bisulfit-Konversionseffizienz. Die Abbildung stellt den *Quality Control Bar Plot* (generiert durch Anwendung von RnBeads (Details siehe 2.2.2.2.1 und 2.2.5.1.1)) des Teilprojekts P1 (HLHS-RF-Analyse) beispielhaft von einer Bisulfit-Konversionskontrollsonde (BSK-Sonde) des Infinium I-Assays dar. Auf der y-Achse ist die Fluoreszenzsignalintensität in FE dargestellt und auf der x-Achse die jeweiligen analysierten Proben des Teilprojekts P1. Bei effizienter Bisulfit-Konversion der analysierten Proben des Teilprojekts P1 wurden bei der hier dargestellten „unkonvertiert“-BSK-Sonde „BS Conversion I U2“ Fluoreszenzsignalintensitäten entsprechend eines Hintergrundsignals (< 200 FE) erwartet. Die Probe HLHS_157 zeigte eine Fluoreszenzsignalintensität > 2000 FE, während die restlichen Proben geringe Signalintensitäten im Bereich des erwarteten Hintergrundsignals (< 200 FE) zeigten. Die Probe HLHS_157 wurde daher aus der Analyse ausgeschlossen.

³³ Proben, die auf dem 450K-Array hybridisiert wurden

Eine weitere Qualitätskontrolle beinhaltete die **Analyse der Korrelationsmatrices** (Details siehe 2.2.2.2.1 und 2.2.4). In den jeweiligen Analysen der Teilprojekte wurden insgesamt vier der 78 Hybridisierungen anhand der Analyse der CpG-Gesamtkorrelation ausgeschlossen. Die Ergebnisse der CpG-Gesamtkorrelationen werden in den einzelnen Teilprojekten näher beschrieben. Die SNP-Korrelation (Abbildung 26) ergab, wie erwartet, eine geringe Korrelation zwischen unterschiedlichen Individuen (blaue bis grünliche Färbung), und eine hohe Korrelation bei Gewebeproben bzw. biologischen Replikaten desselben Individuums mit Korrelationskoeffizienten von $r > 0,999$. Die IAS-Gewebeproben der monozygoten Zwillinge HLHS_22 und HLHS_21 (Nr. 46/47 in Abbildung 26) ergaben hohe Korrelationskoeffizienten von $r > 0,9994$. Die Analyse der SNP-Korrelation ergab keinen Probenausschluss. Die Regressionsanalyse von technischen Duplikaten der Proben F_FetalHerz und M_Fetalherz sowie von biologischen Replikaten aus LV-, RV-, LA- und RA-Gewebe der Patientenproben HLHS-TGA_117 und Lvan_126 ergab jeweils eine hohe Korrelation der DNA-Methylierungswerte mit $R^2 > 0,99$ (siehe Abschnitt 4.1.3.2 Abbildung 58).

Der Datensatz der 485577 Loci des 450K-Arrays wurde im Anschluss an die Qualitätskontrolle mittels RnBeads *R-Package* normalisiert und die CpG-Loci nach Filterkriterien, wie in Abschnitt 2.2.2.2.1 beschrieben, gefiltert. In Tabelle 27 ist pro Teilprojekt eine **Zusammenfassung der Anzahl ausgeschlossener Hybridisierungen und ausgeschlossener CpG-Loci** (durchschnittlicher Ausschluss von 7,5 % der CpG-Loci nach GS-Filterung (siehe 2.2.2.2.1)) mit Angaben zur Anzahl der final mittels QluCore-Software (Details siehe 2.2.5.2) analysierten CpG-Loci aufgeführt. Hierbei ist zu beachten, dass einzelne Hybridisierungen in mehrere Teilprojekte eingegangen sind und jeweils Teilprojekt-spezifisch analysiert wurden. Insgesamt wurde eine Hybridisierung aufgrund unvollständiger Bisulfitkonversion und vier Hybridisierungen nach Analyse der CpG-Gesamtkorrelation aus allen 78 Hybridisierungen ausgeschlossen, was einem Gesamtausschluss von 6,4 % entsprach.

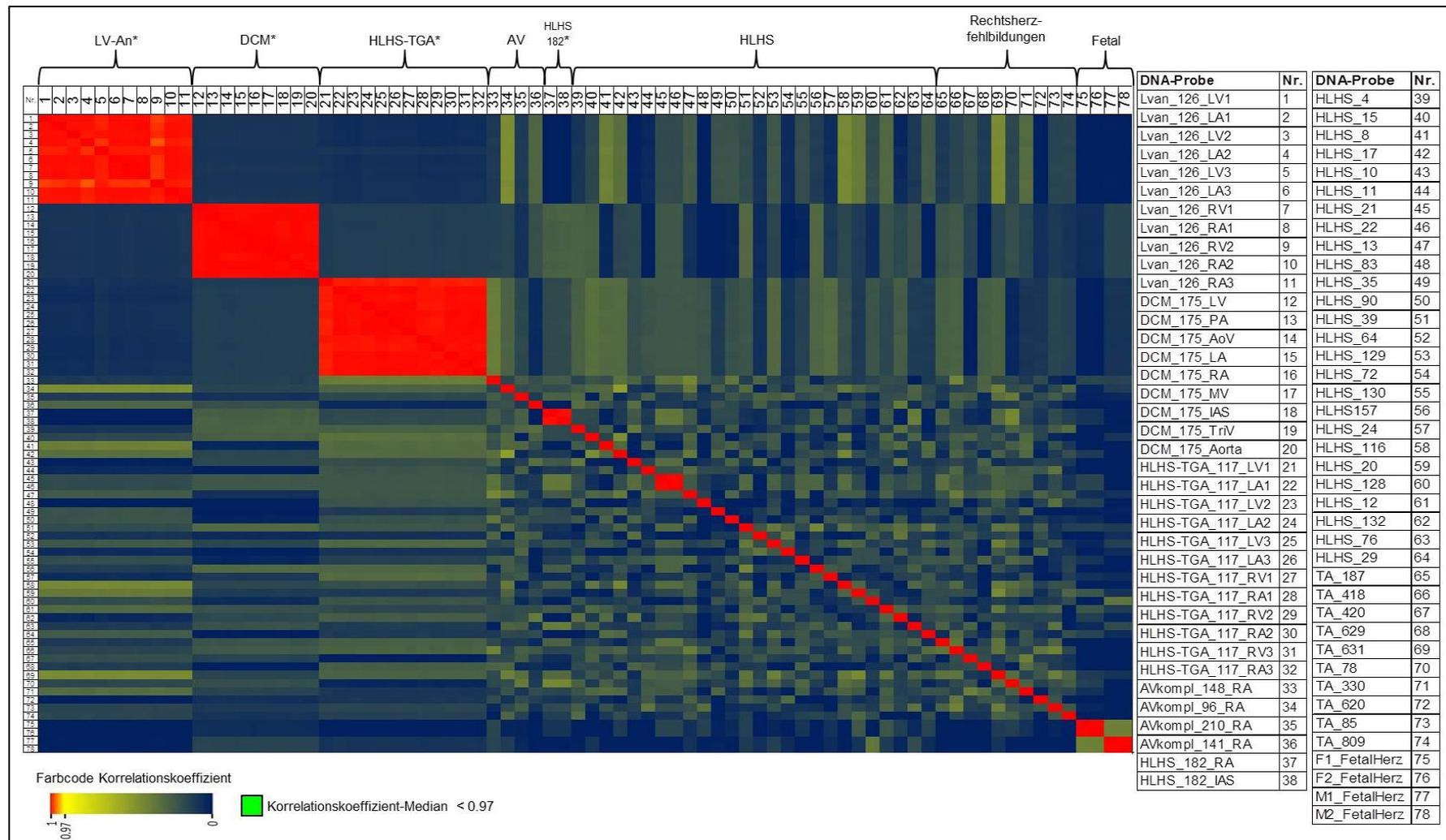


Abbildung 26 SNP-Korrelation (Pearson-Methode) der AVG-Beta-Werte von 78 DNA-Proben verschiedener Herzgewebetypen, die auf einem 450K-Array hybridisiert wurden. Dargestellt sind die Korrelationskoeffizienten-Werte (r) mit einer farblich bedingten Formatierung von niedrigen r-Werten und hohen r-Werten in rot (siehe Farbcode rechts der Abbildung). Jede Spalte (und die gleiche Zeilennummer) steht für eine DNA-Probe, mit einer Nummerierung der Proben von links nach rechts (1 bis 78) der Spalten (bzw. oben nach unten der Zeilen) entsprechend der Tabellen rechts der Korrelationsmatrix. Die jeweiligen Probenbezeichnungen pro Nummer sind in der Tabelle aufgeführt. Die Gewebeprouben desselben Individuums (Probenbezeichnung: Lvan, DCM, HLHS-TGA, HLHS_182, FetalHerz) zeigen eine hohe Korrelation mit $r > 0,999$, was durch eine rote Einfärbung zu erkennen ist. Die IAS-Gewebeprouben der monozygoten Zwillinge HLHS_22 und HLHS_21 (Nr. 46/47) ergaben hohe Korrelationskoeffizienten von $r > 0,9994$.

Tabelle 27 Übersicht über die Anzahl ausgeschlossener Hybridisierungen und ausgeschlossener CpG-Loci, sowie der final analysierten Hybridisierungen und Loci pro Teilprojekt

Kurzbeschreibung Teilprojekt (Kapitel-Abschnitt)	Anzahl Hybrid. ¹ [N]	Anzahl ausgeschlossener Hybridisierungen nach LCR (Median LCR [%])	Anzahl ausgeschlossener Hybridisierungen nach Analyse der Bisulfitkonversions-Effizienz	Anzahl (und Prozentangabe) ausgeschlossener Loci nach GS-Filterung (bezogen auf 485577 Loci des 450K-Arrays)	Anzahl ausgeschlossener Hybridisierungen nach Analyse der Korrelationsmatrix [N]		Anzahl [N] eingeschlossener Loci bzw-Hybridisierungen in der finalen statistischen Analyse	
					CpG-Gesamt ² -Korr.	SNP-Korr.	Anzahl CpG-Loci [N]	Anzahl Hybrid. [N]
P1: HLHS im Vergleich zu Rechtsherzfehlbildungen (3.2.2.1)	36	0 (99,96 %)	1	36125 (7,4 %)	2	0	449452	33
P2: Herz-Mapping - Analyse der Gesamtherz-Methylierung (3.2.2.2)	77 (MW: 60)	0 (99,93 %)	0	37756 (7,8 %)	0	0	447821	60 (MW)
P3: Herz-Mapping: Atrium- und Ventrikelsignatur (3.2.2.3)	68 (MW: 53)	0 (99,95 %)	0	36763 (7,6 %)	4	0	448814	49 (MW)
P4: Muskel- und Gefäßsignatur bei dilatativer Kardiomyopathie (3.2.2.4)	9	0 (99,96 %)	0	35328 (7,3 %)	0	0	450249	9
P5: HLHS im Vergleich zu normalem fetalem Herzgewebe (3.2.2.5)	21 (MW: 19)	0 (99,98 %)	0	35543 (7,3 %)	0	0	450034	19 (MW)

MW: Probenanzahl nach Mittelwertbildung bei Einschluss biologischer Replikate; LCR: *Loci Call Rate*; GS-Filterung: Filterung der CpG-Loci wie beschrieben in Abschnitt 2.2.2.2.1; Korr.: Korrelationsmatrix; Anzahl Hybrid.: Anzahl Hybridisierungen. ¹die Hybridisierungen gingen teilweise in mehr als einem Teilprojekt zur Analyse ein und einige DNA-Proben wurden als biologische Replikate mehrfach hybridisiert, ²Begriff schließt ch-Loci mit ein

3.2.2.1 Alterseffekte der DNA-Methylierung in IAS-Gewebeproben von Fällen mit HLHS und Rechtsherzfehlbildungen

Epigenetische Prozesse stellen einen wichtigen Einflussfaktor in der Herzentwicklung dar [37, 317 - 319]. Um daher die epigenetischen Einflüsse auf die Entwicklung des HLHS zu untersuchen, wurden IAS-Gewebeproben von 26 Patienten mit HLHS (siehe Tabelle 1, P1) hinsichtlich der DNA-Methylierung mittels 450K-Array analysiert. Als Kontrollproben dienten 10 IAS-Proben von Patienten mit Rechtsherzfehlbildungen (Beschreibung des Patientenkollektivs siehe 2.1.1.1, Abbildung 5), sodass insgesamt 36 Hybridisierungen analysiert wurden. Zwar ließen sich in diesem Teilprojekt keine signifikant differenziell methylierten CpG-Loci beim Vergleich des DNA-Methylierungsmusters des Herzgewebes von HLHS-Patienten und Patienten mit Rechtsherzfehlbildungen identifizieren, jedoch konnten Alterseffekte der DNA-Methylierung im IAS-Gewebe detektiert werden.

3.2.2.1.1 Qualitätskontrolle

Die Analyse der Bisulfit-Konversionseffizienz mittels RnBeads-Anwendung (siehe 2.2.2.2.1 und 2.2.5.1.1)³⁴ ergab den Probenausschluss der Probe HLHS_157 (siehe 3.2.2, Abbildung 25). Die CpG-Gesamtkorrelation (Details siehe 2.2.4)³⁴ der AVG-Beta-Werte der nun insgesamt 35 Hybridisierungen resultierte bei allen IAS-Gewebeproben in hohen r-Werten nahe $r = 1$, mit Ausnahme der Proben HLHS_11 und TA_629 (Abbildung 27). Mit einem Median über alle Korrelationskoeffizienten kleiner als der Grenzwert von $r \geq 0,97$ (Details siehe 2.2.4)³⁴ wurden die Proben HLHS_11 ($r = 0,96$) und TA_629 ($r = 0,94$) aus der finalen statistischen Analyse ausgeschlossen. Die Intragruppen-Heterogenität konnte somit für den nachfolgenden T-Test verringert werden. Insgesamt wurden nach der Qualitätsfilterung drei Hybridisierungen exkludiert, sodass final 24 HLHS-Proben und neun Proben von Patienten mit Rechtsherzfehlbildungen analysiert wurden. Die Normalisierung und GS-Filterung (siehe 2.2.2.2.1)³⁴ der 450K-Rohdaten der nun insgesamt 33 IAS-Gewebeproben ergab einen Ausschluss von 36125 CpG-Loci und 449452 CpG-Loci gingen in die weitere statistische Analyse ein (Tabelle 27, P1).

³⁴ Die Qualitätskontrolle der nachfolgenden Teilprojekte P2 bis P5 erfolgte nach demselben Schema, sodass die Verweise zu den Methoden der Qualitätskontrolle in diesen Teilprojekten im Text nicht explizit erwähnt werden.

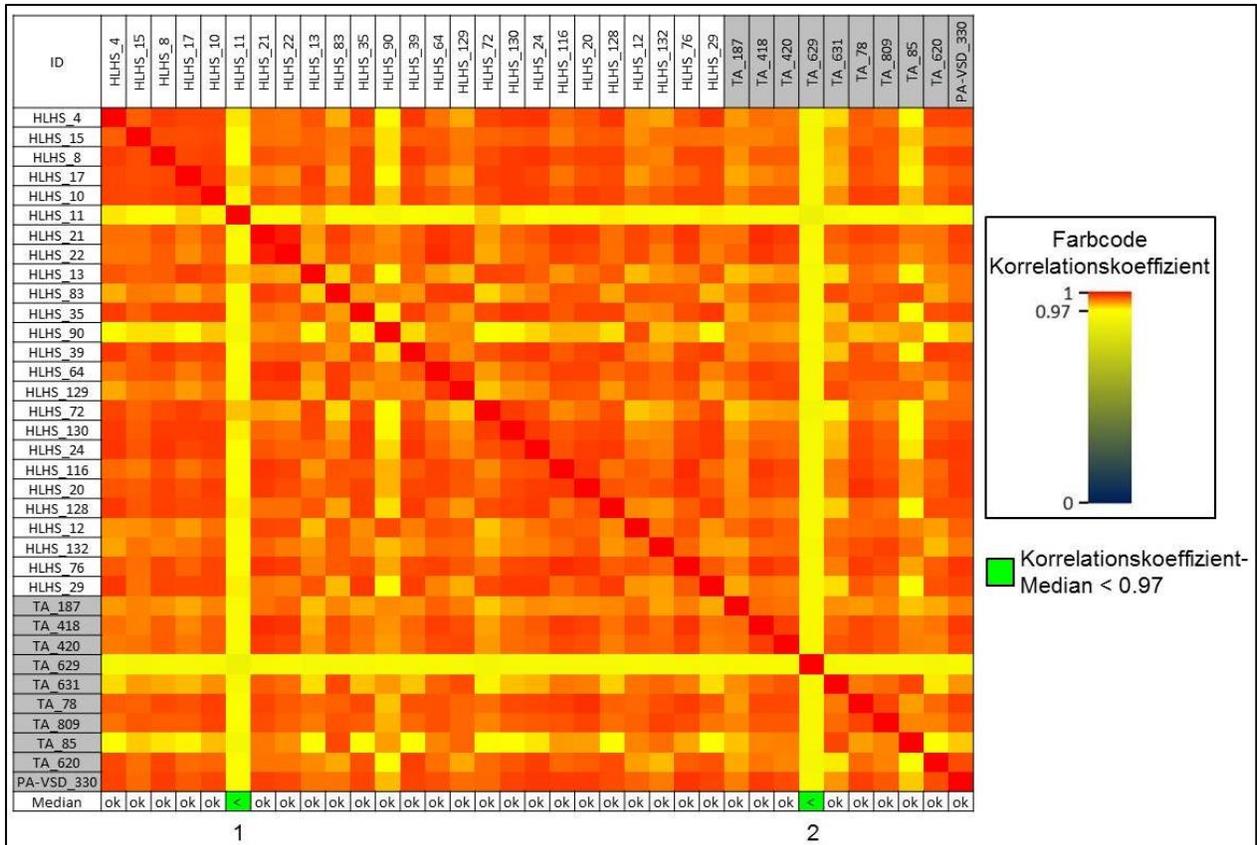


Abbildung 27 CpG-Gesamtkorrelation der AVG-Beta-Werte von IAS-Gewebeproben, hybridisiert auf einem 450K-Array, von 25 Patienten mit HLHS und 10 Patienten mit Rechtsherzfehlbildungen. Dargestellt sind die Korrelationskoeffizient-Werte (r) nach Pearson-Methode mit einer farblich bedingten Formatierung von niedrigen r -Werten in blau und hohen r -Werten in rot (siehe Farbcode). Jede Spalte (und die gleiche Zeilennummer) steht für eine DNA-Probe, mit den HLHS-Proben in weiß und Rechtsherzfehlbildungen-Proben in grau oberhalb und links der Korrelationsmatrix. Unterhalb der Korrelationsmatrix ist der Median über alle Korrelationskoeffizienten der jeweiligen Spalte/DNA-Probe dargestellt (Median $> 0,97$ = „ok“, Median $< 0,97$ = „<“). In grün hinterlegt sind Korrelationskoeffizienten von $r < 0,97$. Die Hybridisierungen 1: HLHS_11 ($r = 0,96$) und 2: TA_629 ($r = 0,94$) wurden aus der finalen statistischen Analyse ausgeschlossen.

3.2.2.1.2 Identifizierung von 11 Genen mit einem altersabhängigen DNA-Methylierungsprofil im IAS-Gewebe von HLHS-Patienten und Patienten mit Rechtsherzfehlbildungen

Anhand der 449452 CpG-Loci wurde zunächst eine **unsupervidierte Analyse** (Varianzfilter $\sigma/\sigma_{\max} \geq 0,4$) durchgeführt, um zu untersuchen, ob eine Segregation der Gruppen HLHS und Rechtsherzfehlbildungen festzustellen ist. Die Hauptkomponentenanalyse (*Principal Component Analysis*, PCA) der resultierenden 1205 CpGs mit $\sigma/\sigma_{\max} \geq 0,4$ ergab hierbei keine

Segregation der IAS-Gewebeproben von HLHS-Patienten und Patienten mit Rechtsherzfehlbildungen (Abbildung 28 A).

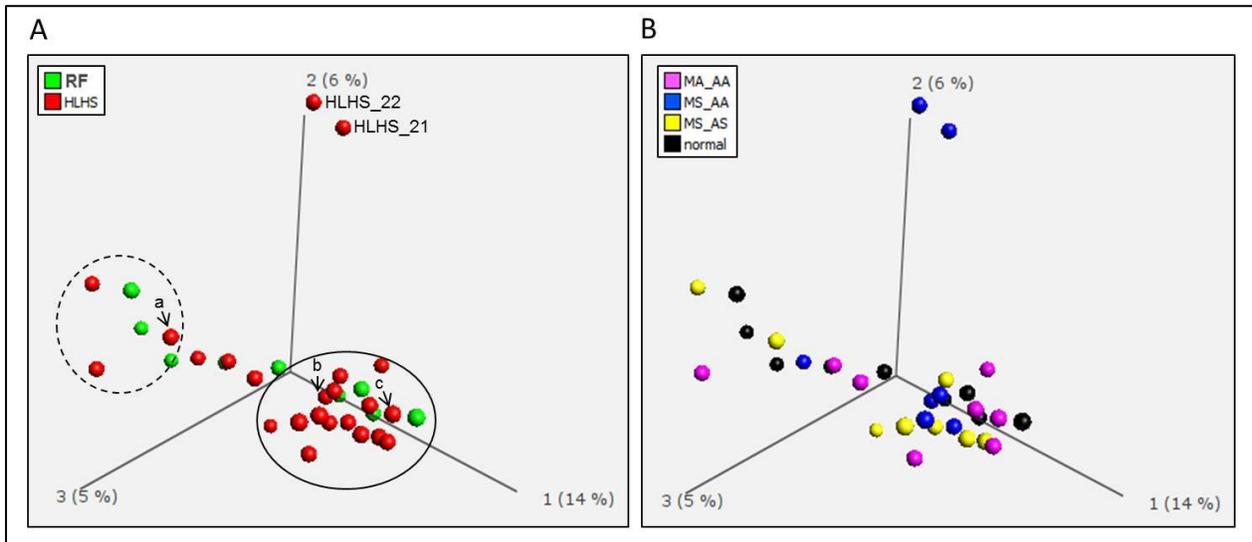


Abbildung 28 Unsupervidierte Analyse als PCA-Darstellung (Varianzfilter $\sigma/\sigma_{\max} \geq 0,4$) der AVG-Beta-Werte von IAS-Gewebeproben von 24 Patienten mit HLHS und neun Patienten mit Rechtsherzfehlbildungen, die mittels 450K-Array analysiert wurden. Dargestellt ist die Segregation der IAS-Gewebeproben basierend auf 1205/449452 CpGs mit $\sigma/\sigma_{\max} \geq 0,4$. A: Die Gruppen HLHS (rote Kugeln) und Rechtsherzfehlbildungen (RF) (grüne Kugeln) zeigten keine Trennung in unterschiedliche Gruppierungen. Der Kreis mit gestrichelter Linie umfasst IAS-Gewebeproben mit hohem mAG von 2414,5 d (Ausnahme a: HLHS_12 4 d) und der Kreis mit solider Linie umfasst IAS-Gewebeproben mit niedrigem mAG von 5,4 d (Ausnahme b: HLHS_116 44 d, c: HLHS_29 106 d). B: Die verschiedenen HLHS-Typen (siehe 2.1.1) MA_AA (Mitralatresie/Aortenatresie, rosa Kugeln), MS_AA (Mitralstenose/Aortenatresie, blaue Kugeln), MS_AS (Mitralstenose/Aortenstenose, gelbe Kugeln) zeigten keine Trennung in unterschiedliche Gruppierungen. Die Proben der Patienten mit Rechtsherzfehlbildungen („normal“) sind als schwarze Kugeln dargestellt. Normalisierung zur graphischen Darstellung: $m = 0$.

Tendenziell zeigte sich jedoch eine Segregation von IAS-Gewebeproben mit unterschiedlichem Patientenalter bei Gewebeentnahme (AG). Hierbei waren IAS-Gewebeproben mit hohem AG (mittleres (m)AG = 2414,5 d) zusammen gruppiert (Abbildung 28 A, gestrichelter Kreis). Auch „junge“³⁵ IAS-Gewebeproben mit niedrigem mAG (5,4 d) bildeten eine gemeinsame Gruppierung (Abbildung 28 A, solider Kreis). Ausnahmen bildeten lediglich die „junge“ Probe HLHS_12 (4 d; Abbildung 28 A „a“), die mit den „alten“ Proben gruppierte und die „alten“ Proben HLHS_116 (44 d; Abbildung 28 A „b“) und HLHS_29 (106 d; Abbildung 28 A „c“), die mit den „jungen“ Proben gruppierten. Die IAS-Gewebeproben der monozygoten Zwillinge HLHS_21 und HLHS_22 segregierten gesondert von den restlichen Proben, wobei hier der individuelle Effekt durch die hohe Ähnlichkeit der AVG-Beta-Werte der beiden Proben eine Rolle spielt. Betrachtete man die unterschiedlichen HLHS-Typen (siehe 2.1.1) Mitralstenose/Aortenstenose (MS/AS),

³⁵ Begriffe „junge“, mittlere“, „alte“ Gewebeprobe bezieht sich grundsätzlich auf das Alter des Patienten bei Gewebeentnahme entsprechend der Alterskategorien aus 2.1.1.1 (Tabelle 1).

Mitralstenose/Aortenatresie (MS/MA) und Mitralatresie/Aortenatresie (MA/AA), so zeigten diese keine Trennung in unterschiedliche Gruppen (Abbildung 28 B). Die HLHS-Typen ließen sich somit nicht durch eine differenzielle DNA-Methylierung unterscheiden.

Eine supervidierte Analyse in Form eines **T-Tests** wurde zum Vergleich der beiden Gruppen HLHS und Rechtsherzfehlbildungen durchgeführt, um zu analysieren, ob eine differenzielle DNA-Methylierung zwischen IAS-Gewebeproben von HLHS-Patienten und dem Kontrollkollektiv aus Rechtsherzfehlbildungen vorlag. Basierend auf dem Varianzfilter $\sigma/\sigma_{\max} \geq 0,4$ ergaben sich bei einem q-Wert (*False Discovery Rate* (FDR)) von $q \leq 0,05$ ³⁶ keine CpG-Loci, die signifikant differenziell methyliert zwischen den Gruppen HLHS und Rechtsherzfehlbildungen vorlagen. Auch eine Lockerung der statistischen Parameter durch Verringerung des Varianzfilters auf $\sigma/\sigma_{\max} \geq 0,2$ ergab bei $q \leq 0,05$ keine CpG-Loci, die dies erfüllten. Somit konnten keine signifikant differenziell methylierten CpG-Loci in IAS-Gewebeproben von Patienten mit HLHS im Vergleich zu Patienten mit Rechtsherzfehlbildungen detektiert werden.

Um den Einfluss des Patientenalters bei Gewebeentnahme zu analysieren, welches in der unsupervidierten Analyse tendenziell eine unterschiedliche Gruppierung von „jungen“ und „alten“ IAS-Gewebeproben (PCA Abbildung 28 A) zeigte, wurde eine **Rangregressions-Analyse** (siehe 2.2.4 und 2.2.5.2) anhand der Gewebealter-Daten durchgeführt. Bei einem q-Wert $\leq 3 \times 10^{-4}$ ($\sigma/\sigma_{\max} \geq 0,2$) lagen 1104 CpG-Loci bei den 33 IAS-Gewebeproben altersabhängig signifikant differenziell methyliert vor. In einer Heatmap-Darstellung wurden die IAS-Gewebeproben nach absolutem Alter sortiert (Abbildung 29, von links nach rechts „junge“ bis „alte“ Proben) und die 1104 CpG-Loci hierarchisch geclustert. Hierbei zeigte sich tendenziell eine Hypomethylierung (niedrige AVG-Beta-Werte) von ca. 80 % der CpG-Loci bei „jungen“ IAS-Gewebeproben (mAG ≤ 3 Wochen) im Vergleich zu „alten“ Gewebeproben (mAG > 30 Wochen), die bei diesen CpG-Loci hypermethyliert (hohe AVG-Beta-Werte) vorlagen. Betrachtete man IAS-Gewebeproben „mittleren“ Alters, so zeigten diese tendenziell eine „Hemimethylierung“ (AVG-Beta-Werte um 0,5). Anhand der Heatmap-Darstellung wurden Alterskategorien erstellt mit folgenden Kategorien: „jung“: ≤ 3 Wochen, „mittel“: > 3 Wochen bis ≤ 30 Wochen und „alt“: > 30 Wochen. Diese wurden nachträglich unter der Bezeichnung „Alter_Kategorien“ in die Heatmap-Darstellung aufgenommen (Abbildung 29). Drei IAS-Gewebeproben – HLHS_12, HLHS_29 und HLHS_116 – zeigten hierbei jedoch besonders auffällig ein abweichendes Methylierungsmuster im Kontext der jeweiligen Alterskategorie, wie bereits in der unsupervidierten Analyse (Abbildung 28 A) zu erkennen war. Insgesamt gesehen war jedoch

³⁶ $q \leq 0,05$ wurde als untere Signifikanzgrenze in dieser Arbeit festgelegt

deutlich die Tendenz der vermehrten Hypomethylierung „junger“ und Hypermethylierung „alter“ IAS-Gewebeproben zu erkennen, sodass diese definierten Alterskategorien bei Analysen weiterer Teilprojekte übernommen wurden.

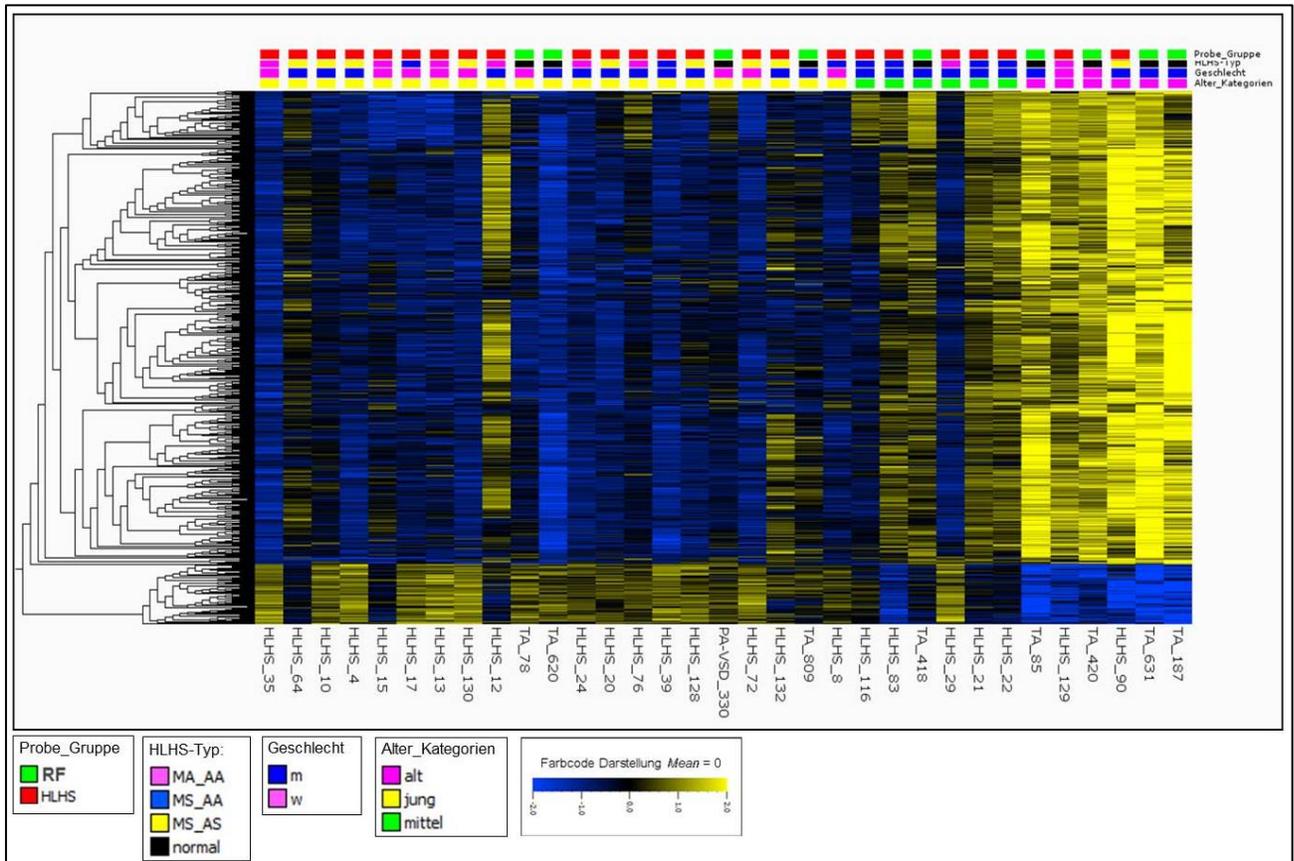


Abbildung 29 Heatmap der Rangregressions-Analyse anhand des Alters bei Gewebeentnahme der AVG-Beta-Werte von IAS-Gewebeproben von 24 Patienten mit HLHS und neun Patienten mit Rechtsherzfehlbildungen, die mittels 450K-Array analysiert wurden. Bei $\sigma/\sigma_{\max} \geq 0,2$ und $q \leq 3 \times 10^{-4}$ lagen 1104/449452 CpG-Loci bei insgesamt 33 IAS-Gewebeproben altersabhängig differenziell methyliert vor. Jede Zeile stellt einen CpG-Locus dar, mit niedrigen AVG-Beta-Werten in blau und hohen AVG-Beta-Werten in gelb (siehe Farbcode). Jede Spalte stellt eine DNA-Probe dar, mit den Bezeichnungen der einzelnen Proben unterhalb der Heatmap. Oberhalb der Heatmap sind pro Probe entsprechend der Farbcode-Legenden folgende Kategorien farblich dargestellt (von oben nach unten): die Probengruppe (RF = Rechtsherzfehlbildung), der HLHS-Typ (RF-Proben als „normal“ bezeichnet), das Geschlecht, und die Alterskategorien („jung“ (≤ 3 Wochen), „mittel“ (> 3 Wochen bis ≤ 30 Wochen) und „alt“ (ab > 30 Wochen)). Normalisierung zur graphischen Darstellung: $m = 0$.

Die PCA-Darstellung (Abbildung 30) der Rangregressionsanalyse der 1104 DM-CpG-Loci verdeutlichte die Kategorien der drei Gruppierungen „jung“, „mittel“ und „alt“. Das abweichende Methylierungsmuster im Kontext der jeweiligen Alterskategorie der Proben HLHS_12, HLHS_29 und HLHS_116 ist in Abbildung 30 deutlich erkennbar. HLHS_12 mit 4 d (Kategorie „jung“) segregierte mit „mittel“ bis „alten“ Proben, während HLHS_29 (106 d) und HLHS_116 (44 d) der Kategorie „mittel“ tendenziell mit „jungen“ Proben in einer Gruppierung auftraten (Abbildung 30).

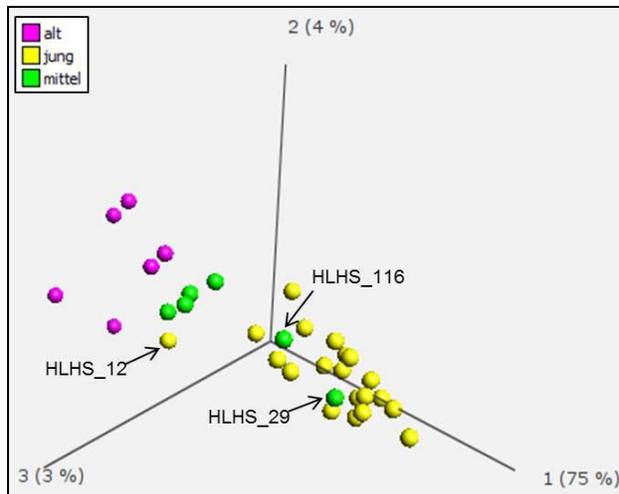


Abbildung 30 PCA-Darstellung der Rangregressions-Analyse anhand des Alters bei Gewebeentnahme von IAS-Gewebeproben von 24 Patienten mit HLHS und neun Patienten mit Rechtsherzfehlbildungen, die mittels 450K-Array analysiert wurden. Dargestellt ist die Segregation der IAS-Gewebeproben basierend auf altersabhängig differenziell methylierten 1205 CpG-Loci der Rangregressions-Analyse anhand Gewebealter-Daten mit $\sigma/\sigma_{\max} \geq 0,2$ und $q \leq 3 \times 10^{-4}$. Die visuell anhand der Heatmap (Abbildung 29) definierten Alterskategorien („jung“ (≤ 3 Wochen) in gelb, „mittel“ (> 3 Wochen bis ≤ 30 Wochen) in grün und „alt“ (ab > 30 Wochen) in rosa ließen sich tendenziell als segregierende Gruppierungen erkennen. Ausnahmen bildeten HLHS_12 (4 d), HLHS_29 (106 d) und HLHS_116 (44 d). Normalisierung zur graphischen Darstellung: $m = 0$.

Es konnte keine signifikante Anreicherung von Genen für bestimmte *Pathways* oder biologische Prozesse in einer **Anreicherungsanalyse** mittels WebGestalt-Tool (siehe 2.2.5.3) innerhalb der assoziierten Gene der altersabhängig differenziell methylierten 1205 CpG-Loci festgestellt werden. Betrachtete man die Gene mit einer besonders hohen Anzahl von DM-CpG-Loci pro Gen (hier ≥ 4 DM-CpG-Loci), so ließen sich 11 Gene – *ATP11A*, *C10orf140*, *DENND3*, *DNMT3A*, *EBF4*, *ERI3*, *ESR1*, *FBN2*, *MAD1L1*, *SNED1* und *TMEM140* – mit einem möglichen altersabhängigen DNA-Methylierungsprofil im IAS-Gewebe identifizieren. Das Gen *ESR1*, welches zuvor schon in der Literatur mit einer altersabhängigen differenziellen DNA-Methylierung detektiert werden konnte [146, 149], zeigte bei 4/5 DM-CpG-Loci eine Hypermethylierung bei „alten“ Gewebeproben (AG > 30 Wochen) im Vergleich zu „jungen“ IAS-Gewebeproben (AG ≤ 3 Wochen). Um altersspezifische Effekte zu vermeiden, wurden diese Gene von den weiteren Teilprojektanalysen ausgeschlossen.

Zusammenfassend konnte keine differenzielle DNA-Methylierung im IAS-Gewebe von Patienten mit HLHS im Vergleich zum Kontrollkollektiv aus Patienten mit Rechtsherzfehlbildungen detektiert werden. Es konnte jedoch ein altersabhängiges DNA-Methylierungsmuster mit einer Hypomethylierung bei „jungen“ IAS-Gewebeproben (AG ≤ 3 Wochen) im Vergleich zu „alten“ Gewebeproben (AG > 30 Wochen), unabhängig von der kardialen Diagnose, festgestellt werden.

3.2.2.2 Verschiedene Regionen des Herzens lassen sich durch differenzielle DNA-Methylierungsprofile unterscheiden

Nachdem der Vergleich der DNA-Methylierung von IAS-Gewebeproben von HLHS-Patienten und der Kontrollkohorte bestehend aus Patienten mit Rechtsherzfehlbildungen (siehe 3.2.2.1) keine signifikant differenziell methylierten CpG-Loci ergeben hatte, erfolgte in diesem Teilprojekt P2 zunächst ein systematischer Vergleich des DNA-Methylierungsprofils verschiedener Herzgewebetypen des gesamten Herzens (Herz-Mapping). Die Fragestellung, ob verschiedene kardiale Regionen sich durch eine differenzielle DNA-Methylierung unterscheiden lassen, sollte hiermit bearbeitet werden. Zusätzlich wurde fetales Herzgewebe gesunder Individuen (BioCat, (Heidelberg, Deutschland)) auf dem 450K-Array hybridisiert, um in einer nachfolgenden Analyse (siehe Teilprojekt P5, 3.2.2.5) die DNA-Methylierung bei Patienten mit HLHS im Vergleich zu gesunden Normalkontrollen analysieren zu können. Zur Einordnung des Herzgewebetyps der fetalen Herzgewebeproben, die nach Angabe des Anbieters (Biocat (Heidelberg, Deutschland)) aus dem „Gesamtherz“ stammten, wurde in diesem Teilprojekt P2 das Methylierungsprofil der fetalen Herzgewebeproben mit dem von Herzgewebeproben verschiedener kardialer Regionen verglichen.

Die Herz-Mapping-Analyse umfasste folgende Herzgewebeproben von Patienten mit verschiedenen CHDs, die mittels 450K-Array (siehe 2.2.2.2) analysiert wurden:

- Hybridisierungen aus Teilprojekt P1 (35 IAS-Proben (ohne HLHS_157), siehe 3.2.2.1)
- HLHS_182_IAS, HLHS_182_RA
- 3xLV, 3xRV, 3xLA, 3xRA von einem Patienten mit HLHS-TGA (biologische Replikate)
- 3xLV, 2xRV, 3xLA, 3xRA von einem Patienten mit LV-An (biologische Replikate)
- IAS, LA, RA, LV, MV, TV, PA, Aorta, AoV von einem Patienten mit DCM
- RA-Gewebe von vier Patienten mit AVSD
- fetales Herzkontrollgewebe (s.o.) von einem gesunden weiblichen Donor (40. Woche) und einem gesunden männlichen Donor (25. Woche), je als Duplikate

Insgesamt wurden 77 Hybridisierungen verschiedener Herzgewebetypen in der Herz-Mapping-Analyse untersucht.

3.2.2.2.1 Qualitätskontrolle

Um zu untersuchen, ob sich die verschiedenen kardialen Regionen mithilfe der Erhebung von Korrelationskoeffizienten (r) unterscheiden lassen, wurde eine CpG-Gesamtkorrelation der AVG-Beta-Werte der 77 Hybridisierungen durchgeführt. Hierbei sollte die Frage beantwortet werden, ob einerseits Proben innerhalb derselben Gewebegruppe³⁷ durch homogene r -Werte erkennbar sind, und ob andererseits Proben unterschiedlicher Gewebegruppen sich durch differenzielle r -Werte unterscheiden lassen. Gewebeproben derselben derselben Gewebegruppe, d.h. jeweils atriale (LA, RA) und ventrikuläre Gewebeproben (LV, RV) eines Individuums, zeigten hierbei in der CpG-Gesamtkorrelation homogen hohe r -Mittelwerte von $r > 0,995$ (Abbildung 31, Markierung „A“ und „V“ mit jeweils roter Einfärbung erkennbar). Innerhalb der DCM-Proben zeigten die DNA-Proben von Gefäßen (PA, Aorta) und die „gefäßähnliche“ AoV-Probe eine deutlich abweichende Korrelationskoeffizient-Signatur³⁸ (Abbildung 31, Nr. 1-3 (r -Werte $< 0,97$), Markierung „G“) im Vergleich zu den restlichen Herzgewebeproben des DCM-Patienten. Die Probe AoV wurde bei der Gewebeexzision aus dem Aortenklappenring (Beginn der Aorta *ascendens* [248, 320]) entnommen (siehe 2.1.1.1) und daher in dieser Arbeit im Folgenden als „gefäßähnlich“ bezeichnet. Die Probe AVkompl_210_RA zeigte innerhalb der Gruppe der AVSD-Patienten eine abweichende Korrelationskoeffizient-Signatur (Abbildung 31, Nr. 4 (r -Wert $< 0,97$)). Auch die Proben HLHS_11 und TA_629 zeigten eine abweichende Korrelationskoeffizient-Signatur im Vergleich zu den anderen Proben derselben Diagnosegruppe (Abbildung 31, Nr. 5 bzw. 6 (r -Wert $< 0,97$)). Da in der folgenden Herzmapping-Analyse jedoch die Analyse verschiedener Gewebetypen im Fokus steht, und somit ein Mehrguppenvergleich und kein T-Test durchgeführt werden soll (Details siehe 2.2.4), werden diese Proben nicht aus der Analyse ausgeschlossen. Die technischen Duplikate der fetalen Herzgewebeproben wiesen wie erwartet einen hohen Korrelationskoeffizienten von $r > 0,99$ zueinander auf (Abbildung 31). Die Normalisierung und GS-Filterung der 450K-Rohdaten der insgesamt 77 Herzgewebeproben ergab einen Ausschluss von 37756 CpG-Loci. 447821 CpG-Loci gingen in die finale statistische Analyse ein (Tabelle 27, P2). Von den so prozessierten AVG-Beta-Werten der biologischen Replikate der LV-, RV-, LA-, RA-Proben von LV-An und HLHS-TGA sowie der Duplikate je des weiblichen und männlichen Donors der fetalen Herzgewebeproben wurden dann Mittelwerte gebildet, sodass insgesamt 60 Proben in die finale statistische Analyse gingen.

³⁷ Der Begriff „Gewebegruppe“ umfasst in dieser Arbeit die Einteilung des Herzens in „Atrium“ und „Ventrikel“ sowie in die Kategorien muskuläres Herzgewebe und Gefäß-/gefäß-ähnliches Gewebe.

³⁸ Der Begriff „Korrelationskoeffizient-Signatur“ beschreibt in dieser Arbeit die Signatur von homogenen, hohen r -Werten (r nahe 1, rote Einfärbung in der Korrelationsmatrix) innerhalb einer Gewebegruppe im Zusammenhang mit vergleichsweise dazu niedrigeren r -Werten (orange bis gelbe Färbung in der Korrelationsmatrix) bei Proben unterschiedlicher Gewebegruppen.

3.2.2.2 Identifizierung distinkter DNA-Methylierungsprofile in atrialem und ventrikulärem Herzgewebe sowie in Gefäßstrukturen

Eine **unsupervidierte Analyse** der AVG-Beta-Werte der 60 Herzgewebeproben³⁹ verschiedener kardialer Regionen ergab mit $\sigma/\sigma_{\max} \geq 0,4$ 7763 differenziell methylierte CpG-Loci (DM-CpGs). Die PCA (Abbildung 32) ließ tendenziell je eine Gruppierung atrialer Gewebeproben (LA, RA, IAS), ventrikulärer/„ventrikel-assoziiertes“⁴⁰ Gewebeproben (LV, RV, MV, TV), sowie Gefäß/„gefäß-ähnlicher“ Proben (PA, Aorta, AoV) erkennen. Lediglich die atrialen Proben DCM_175_RA, AVkompl_210_RA und TA_629 (r -Werte $< 0,97$, siehe Abbildung 31) segregierten außerhalb der Gruppierungen (Abbildung 32). Die fetalen Herzgewebeproben zeigten tendenziell eine Gruppierung zusammen mit ventrikulären Proben.

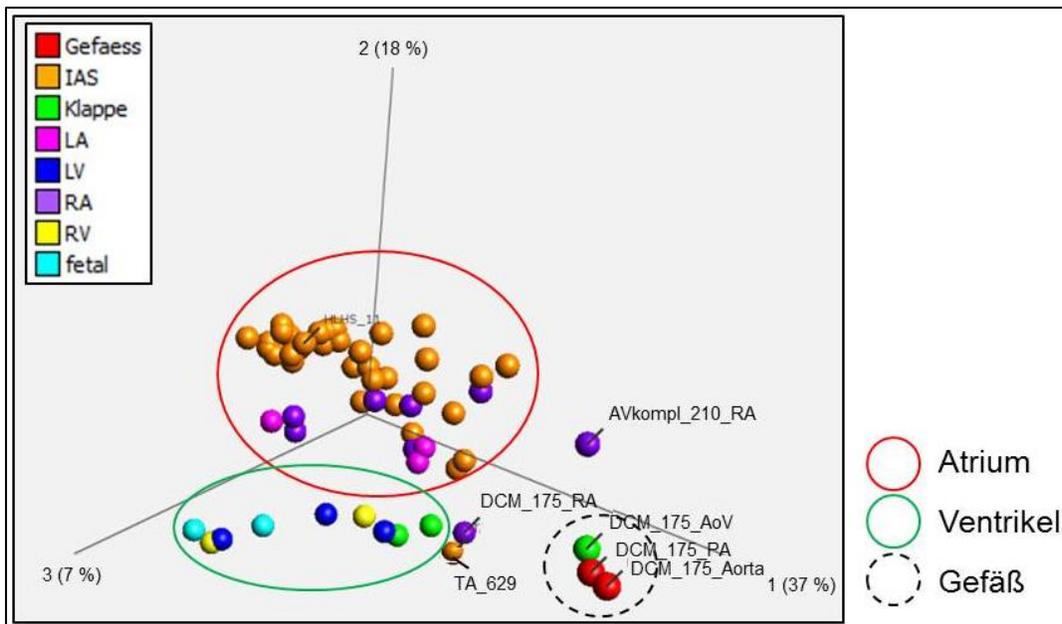


Abbildung 32 Unsupervidierte Analyse als PCA-Darstellung (Varianzfilter $\sigma/\sigma_{\max} \geq 0,4$) der AVG-Beta-Werte von 60 Herzgewebeproben verschiedener kardialer Regionen, die mittels 450K-Array analysiert wurden. Dargestellt ist die Segregation der 60 Herzgewebeproben basierend auf 7763/447821 CpG-Loci mit $\sigma/\sigma_{\max} \geq 0,4$. Farbcode Herzgewebetyp: rot = Gefäß (PA, Aorta), orange = IAS, grün = Herzklappen (MV, TV, AoV), rosa = LA, blau = LV, violett = RA, gelb = RV und türkis = fetales Herzgewebe. Tendenziell war eine Gruppierung von atrialen Gewebeproben (LA, RA, IAS (rote Umrandung)), ventrikulären/„ventrikel-assoziierten“ Gewebeproben (LV, RV, MV, TV (grüne Umrandung)) sowie von Gefäß/„gefäß-ähnlichen“ Proben (PA, Aorta, AoV (gestrichelte Umrandung)) zu erkennen. Proben mit $r < 0,97$ (Abbildung 31) sind in der PCA gesondert beschriftet: DCM_175_RA, DCM_175_PA, DCM_175_Aorta, DCM_175_AoV, AVkompl_210_RA, HLHS_11, TA_629. Normalisierung zur graphischen Darstellung: $m = 0$.

³⁹ Der Begriff schließt berechnete Mittelwerte der AVG-Beta-Werte von Replikaten zusammengefasst als „eine Probe“ mit ein.

⁴⁰ Der Begriff „ventrikel-assoziiert“ umfasst alle Proben, die zusammen mit den Ventrikel-Proben (LV, RV) in der PCA und der Clusteranalyse segregierten.

Die **Varianzanalyse (ANOVA)** (siehe 2.2.5.2) wurde als supervidierte Analyse der definierten Gewebetypen LV, RV, LA, RA, IAS, Gefäß (PA, Aorta), Herzklappen (MV, TV, AoV) und des fetalen Herzgewebes („fetal“) durchgeführt. Die ANOVA ergab bei $q \leq 1 \times 10^{-6}$ ($\sigma/\sigma_{\max} \geq 0,4$) 1404 signifikant differenziell methylierte CpG-Loci zwischen den verschiedenen Gewebetypen (Abbildung 33 A). Auch hier war, wie schon in der unsupervidierten Analyse erkennbar (Abbildung 32), eine Gruppierung atrialer, ventrikulärer/„ventrikel-assoziiertes“ und Gefäß/„gefäß-ähnlicher“ Proben festzustellen. Eine abweichende Segregation im Kontext der Gruppierungen zeigten wiederum DCM_175_RA, AVkompl_210_RA und TA_629 (r -Werte $< 0,97$, siehe Abbildung 31). Schloss man alle DNA-Proben mit einem $r < 0,97$ aus der ANOVA-Analyse aus (siehe beschriftete Proben in Abbildung 33 A), so war eine distinkte Segregation von zwei verschiedenen Gewebegruppen „Atrium“ und „Ventrikel“ erkennbar (Abbildung 33 B, 654 DM-CpGs). Die fetalen Herzgewebeproben gruppieren hierbei wiederum mit den ventrikulären Proben. Im nachfolgenden Teilprojekt P3 (siehe 3.2.2.3) wurde diese Atrium- und Ventrikel-Signatur genauer analysiert.

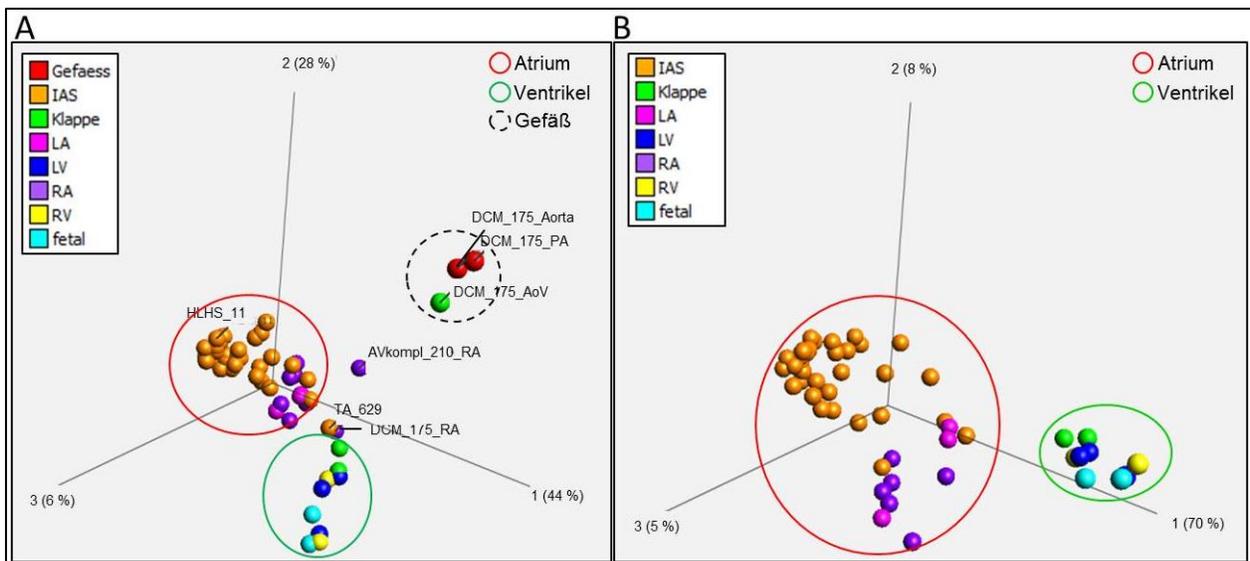


Abbildung 33 PCA-Darstellung der AVG-Beta-Werte der ANOVA der definierten Gewebetypen LV, RV, LA, RA, IAS, Gefäß (PA, Aorta), Herzklappen (MV, TV, AoV) und fetalem Herzgewebe („fetal“), die mittels 450K-Array analysiert wurden. Dargestellt ist die Segregation der verschiedenen Herzgewebeproben basierend auf (A) 1404/447821 CpG-Loci der ANOVA nach Gewebetyp mit $\sigma/\sigma_{\max} \geq 0,4$ und $q \leq 1 \times 10^{-6}$ der 60 Herzgewebeproben und (B) 654/447821 CpG-Loci mit $\sigma/\sigma_{\max} \geq 0,4$ und $q \leq 1 \times 10^{-6}$ der 53 Herzgewebeproben bei Nicht-Einberechnung von sieben Proben mit $r > 0,97$ (DCM_175_RA, DCM_175_PA, DCM_175_Aorta, DCM_175_AoV, AVkompl_210_RA, HLHS 11, TA_629). Proben mit $r > 0,97$ sind in Abbildung A gesondert beschriftet. Farbcode Herzgewebetyp in Abbildung B: rot = Gefäß (PA, Aorta), orange = IAS, grün = Herzklappen (MV, TV, AoV), rosa = LA, blau = LV, violett = RA, gelb = RV und türkis = fetales Herzgewebe. A: Eine Gruppierung von atrialen Gewebeproben (LA, RA, IAS (rote Umrandung)), ventrikulären/„ventrikel-assoziierten“ Gewebeproben (LV, RV, MV, TV (grüne Umrandung)) sowie von Gefäß/„gefäß-ähnlichen“ Proben (PA, Aorta, AoV (gestrichelte Umrandung)) war tendenziell zu erkennen. B: Bei Ausschluss von DNA-Proben mit $r < 0,97$ aus der ANOVA-Analyse (siehe beschriftete Proben in Abb. A) war eine Segregation von zwei verschiedenen Gewebegruppen „Atrium“ (rote Umrandung) und „Ventrikel“ (grüne Umrandung) erkennbar. Normalisierung zur graphischen Darstellung: $m = 0$.

Die Heatmap-Darstellung (Abbildung 34) der ANOVA nach Gewebetyp mit 1404 DM-CpG-Loci (Abbildung 34) verdeutlichte die Gruppierungen atrialer, ventrikulärer/„ventrikel-assoziiertes“ Proben und Gefäß/„gefäß-ähnlicher“ Proben. Die atrialen Herzgewebeprouben bildeten in der Clusteranalyse ein separates Cluster des DNA-Methylierungsprofils (Abbildung 34, „Atrium“). Auch die ventrikulären Herzgewebeprouben bildeten ein Cluster (Abbildung 34, „Ventrikel“). Die fetalen Herzgewebeprouben (Abbildung 34, rote Umrandung) zeigten ein Methylierungsmuster, das dem von LV- und RV-Proben sehr ähnelte. Daher wurde von einer ventrikulären Herkunft des fetalen Herzgewebes ausgegangen. Wie schon in der assoziierten PCA erkennbar (Abbildung 33 A), zeigten die Proben AVkompl_210_RA, DCM_175_RA und TA_629 ein abweichendes Methylierungsmuster in getrennten „Clusterästen“ im Kontext der Gruppierungen (Abbildung 34 mit Sternchen markiert). Herzgefäß-Proben (PA, Aorta), sowie die „gefäß-ähnliche“ Probe AoV (entstammt dem Aortenklappenring (siehe 2.1.1.1)) wiesen ein deutlich abweichendes Methylierungsmuster im Vergleich zu den restlichen Herzgewebeprouben auf (Abbildung 34, „Gefäß“). In dem nachfolgenden Teilprojekt P4 (siehe 3.2.2.4) wurde diese Gefäßsignatur an Herzgewebeprouben des DCM-Patienten genauer analysiert. Betrachtete man die 450K-Array-ID (Sentrix ID) in einer PCA-Darstellung, so waren keine Batch-Effekte erkennbar (siehe Abschnitt 8 Abbildung 60), sodass die segregierenden Gruppierungen von Atrium-, Ventrikel- und Gefäßproben als glaubwürdig eingestuft werden konnten.

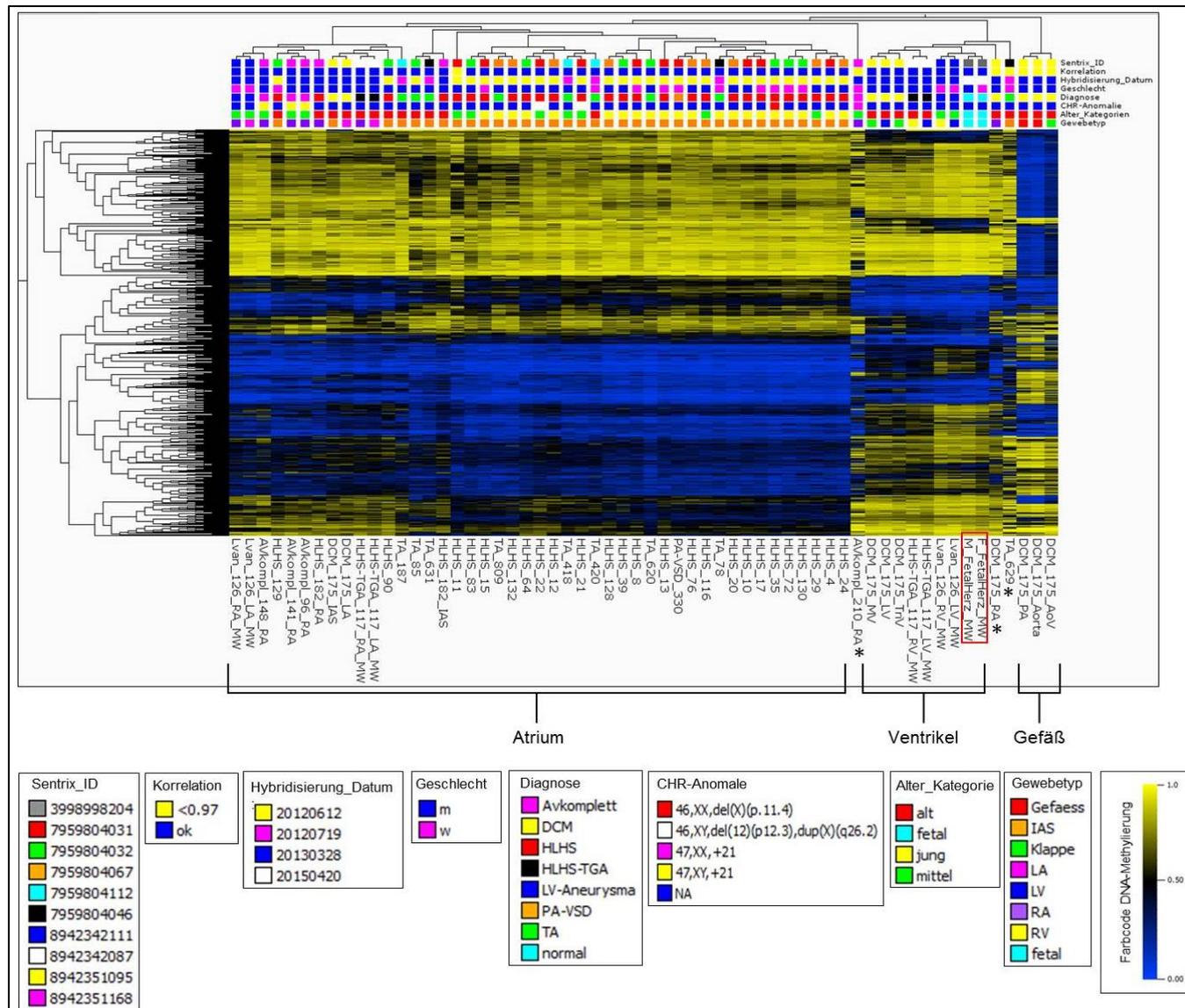


Abbildung 34 Heatmap der AVG-Beta-Werte der ANOVA anhand der definierten Gewebetypen LV, RV, LA, RA, IAS, Gefäß (PA, Aorta), Herzklappen (MV, TV, AoV) und fetalem Herzgewebe („fetal“), die mittels 450K-Array analysiert wurden. Proben und CpG-Loci wurden hierarchisch geclustert. AVG-Beta-Mittelwerte sind hierbei von den Proben Lvan, HLHS-TGA und jeweils der Duplikate von fetalem Herzgewebe (rote Umrandung) des männl. und weibl. Donors eingegangen (siehe Beschriftung „_MW“). Die ANOVA nach Gewebetyp mit $\alpha/\sigma_{max} \geq 0,4$ und $q \leq 1 \times 10^{-6}$ ergab 1404/447821 DM-CpG-Loci zwischen den verschiedenen Herzgewebetypen. Jede Zeile stellt einen CpG-Locus dar, mit niedrigen AVG-Beta-Werten in blau und hohen AVG-Beta-Werten in gelb. Jede Spalte stellt eine DNA-Probe dar, mit den Bezeichnungen der einzelnen Proben unterhalb der Heatmap. Die Gruppierungen Atrium, Ventrikel und Gefäß sind als Klammern zusammengefasst. Proben mit „Gruppierungs-abweichendem“ Methylierungsmuster sind mit einem Sternchen markiert (AVkompl_210_RA, DCM_175_RA, TA_629). Oberhalb der Heatmap sind pro Probe entsprechend der Farbcode-Legenden unterhalb der Heatmap folgende Kategorien farblich dargestellt (von oben nach unten): die 450K-Array-ID (Sentrix_ID), Proben mit $r < 0,97$ (Korrelation), das

Hybridisierungsdatum, das Geschlecht, die kardiale Diagnose, bekannte chromosomale Anomalien (CHR-Anomalie), Alterskategorien („jung“ (≤ 3 Wochen), „mittel“ (> 3 Wochen bis ≤ 30 Wochen) und „alt“ (ab > 30 Wochen)) und Gewebetyp. Normalisierung zur graphischen Darstellung: keine, Einfärbung nach absoluten AVG-Beta-Werten.

Von den 1404 signifikant differenziell methylierten CpG-Loci der ANOVA der verschiedenen Herzgewebetypen konnten 880 CpG-Loci insgesamt 612 Genen⁴¹ zugeordnet werden. Insgesamt acht dieser Gene konnten in der OMIM-Datenbank (Online Mendelian Inheritance in Man [270], siehe 2.1.3.2) in Assoziation mit CHDs gefunden werden (Tabelle 28). Die meisten DM-CpG-Loci pro Gen zeigten: (1) *PRDM16* (24 DM-CpGs), welches zuvor schon bei Fällen mit linksventrikulärer *non-compaction* Kardiomyopathie (LVNC) identifiziert werden konnte [321], (2) *FGFR2* (8 DM-CpGs), identifiziert bei CHD-Fällen mit Apert Syndrom [322] und (3) dem Gen *NR2F2* (5 DM-CpGs), in dem Genvarianten in ursächlichem Zusammenhang bei Fällen mit AVSD und LVOTOs detektiert werden konnten [303]. Aufgrund des Auftretens multipler DM-CpG-Loci pro Gen wurden diese Gene als signifikant in der herzregionspezifischen DNA-Methylierung erachtet.

Tabelle 28 In Assoziation mit CHDs nach OMIM-Datenbank identifizierte Gene mit DM-CpGs in verschiedenen Herzgewebetypen

Gen (HGNC)	Anzahl DM-CpGs	Syndrom (OMIM)	OMIM Nr.	CHD	Referenz
<i>ELN</i>	2	supravalvuläre Aortenstenose	185500	supravalv. AS, PS	(Ewart et al., 1994) [323] (Urbán et al., 2000) [324]
<i>FGFR2</i>	8	Apert Syndrom (APRS)	101200	ASD, VSD, CoA, Dextrokardie	(Slaney et al., 1996) [322]
<i>NR2F2</i>	5	CHD	615779	AVSD, AS, VSD, TOF, HLHS, CoA	(Al Turki et al., 2014) [303]
<i>PRDM16</i>	24	linksventrikuläre <i>non-compaction</i> Kardiomyopathie 8 (LVNC)	615373	LVNC, DCM	(Arndt et al., 2013) [321]
<i>SKI</i>	2	Shprintzen-Goldberg Syndrom (SGS)	182212	MVP, Aortenwurzel-Dilatation	(Doyle et al., 2012) [325]
<i>STRA6</i>	2	Mikrophthalmie Typ 9 (MCOPS9)	601186	ASD, VSD, TOF, PS, CoA	(Pasutto et al., 2007) [326] (Chassaing et al., 2013) [327]
<i>TBX5</i>	3	Holt-Oram Syndrom (HOS)	142900	ASD, VSD, HLHS, PDA	(Basson et al., 1999) [328] (Mori and Bruneau, 2004) [329]
<i>ZEB2</i>	1	Mowat-Wilson Syndrom (MWIS)	235730	ASD, VSD, PS, PDA	(Amiel et al., 2001) [330] (Ishihara et al., 2004) [331]

Supravalv. AS: supravalvuläre Aortenstenose; PS: Pulmonalstenose; TOF: Fallot'sche Tetralogie; CoA: Koarktion der Aorta; LVNC: linksventrikuläre *non-compaction* Kardiomyopathie; MVP: Mitralklappen-Prolaps; PDA: persistierender *Ductus arteriosus*

Insgesamt lag eine Anreicherung von CpG-Loci, die in Assoziation mit RDMRs (reprogrammierungsspezifische DMRs) vorliegen, innerhalb der in unterschiedlichen Herzregionen differenziell methylierten 1404 CpGs vor (*odds ratio* (OR) = 4,65, $p = 2,2 \times 10^{-16}$). Auch Enhancer-assoziierte CpG-Loci zeigten mit OR = 3,25 ($p = 2,2 \times 10^{-16}$) eine signifikante Anreicherung innerhalb der DM-CpGs unterschiedlicher Herzregionen.

⁴¹ Gen-assoziierte CpG-Loci entsprechen CpG-Loci einer "Genregion" (Exons und Introns eines Gens mit Promotorbereich, 5'UTR, und 3'UTR) nach Illumina-Annotation

Um zu analysieren, welche *Pathways* oder biologischen Prozesse eine differenzielle Methylierung zwischen den verschiedenen Herzregionen aufwiesen, wurde eine **Anreicherungsanalyse** mittels WebGestalt-Tool (siehe 2.2.5.3) der 612 Gene mit DM-CpGs durchgeführt. Die GO-Analyse ergab eine signifikante Anreicherung von Genen im übergeordneten biologischen Prozess „developmental process (GO:0032502)“ (232 Gene, $q = 6,99 \times 10^{-12}$), welches das „embryo development (GO:0009790)“ (71 Gene, $q = 5,45 \times 10^{-9}$) mit einbezog.

Zusammenfassend ergab die Analyse der DNA-Methylierung unterschiedlicher Regionen des Herzens distinkte DNA-Methylierungsprofile von atrialen, ventrikulären/“ventrikel-assoziierten“ und Gefäß/“gefäß-ähnlichen“ Proben. Verschiedene Regionen des Herzens lassen sich somit durch eine differenzielle DNA-Methylierung unterscheiden. Die fetalen Herzgewebeproben zeigten in der PCA und der Clusteranalyse ein Ventrikel-ähnliches DNA-Methylierungsmuster, sodass diese als Proben ventrikulärem Ursprungs eingestuft wurden. Insgesamt lag eine Anreicherung von RDMR- und Enhancer-assoziierten CpG-Loci vor, sodass die verschiedenen Herzregionen vornehmlich in regulativen DNA-Regionen eine differenzielle Methylierung zeigten. Einige Gene, wie *PRDM16*, *FGFR2* und *NR2F2*, die nach OMIM-Datenbank in Assoziation mit CHDs identifiziert wurden, wiesen multiple DM-CpG-Loci pro Gen auf. Die herzregionspezifische DNA-Methylierung letztgenannter Gene wurde daher in nachfolgenden Teilprojekten bzw in einer integrativen Analyse nähergehend untersucht.

In den folgenden Teilprojekten P3 (siehe 3.2.2.3) und P4 (siehe 3.2.2.4) werden die mithilfe des Herz-Mappings identifizierten spezifischen Gewebe-Signaturen genauer analysiert.

3.2.2.3 Nachweis einer differenziellen DNA-Methylierungs-Signatur in Atrium- und Ventrikelgewebe

In der vorausgegangenen Herz-Mapping-Analyse des Teilprojekts P2 (siehe 3.2.2.2) konnte bereits in der unsupervidierten Analyse eine separate Gruppierung atrialer und ventrikulärer/“ventrikel-assoziiertes“ Gewebeproben detektiert werden (Abbildung 32 und Abbildung 33 B). In diesem Teilprojekt P3 sollte nun die Fragestellung beantwortet werden, ob sich atriales und ventrikuläres Herzgewebe signifikant hinsichtlich der DNA-Methylierung unterscheiden lassen und welche Gene eine differenzielle DNA-Methylierung aufweisen. Im Folgenden wurden ausschließlich physiologisch klar definierte atriale (definiert als LA, RA, IAS) und ventrikuläre Gewebeproben (definiert als LV, RV) verwendet. Somit wurden die “ventrikel-

assoziierten“ Proben (Herzklappen MV und TV), im Gegensatz zur vorherigen Herz-Mapping-Analyse (siehe 3.2.2.2), nicht in die Analyse einbezogen.

Die Analyse der DNA-Methylierungs-Signatur von Atrium- und Ventrikelgewebe umfasste somit folgende Herzgewebeproben:

- Hybridisierungen aus Teilprojekt P1 (35 IAS-Proben (ohne HLHS_157), siehe 3.2.2.1)
- HLHS_182_IAS, HLHS_182_RA
- 3xLV, 3xRV, 3xLA, 3xRA von einem Patienten mit HLHS-TGA (biologische Replikate)
- 3xLV, 2xRV, 3xLA, 3xRA von einem Patienten mit LV-An (biologische Replikate)
- IAS, LA, RA, LV von einem Patienten mit DCM
- RA-Gewebe von vier Patienten mit AVSD

Insgesamt wurden 68 Hybridisierungen hinsichtlich einer differenziellen DNA-Methylierung zwischen atrialen und ventrikulären Herzgewebeproben analysiert.

3.2.2.3.1 Qualitätskontrolle

Die CpG-Gesamtkorrelation der AVG-Beta-Werte der nun insgesamt 68 Hybridisierungen ergab bei allen IAS-Gewebeproben hohe r-Werte nahe $r = 1$ mit Ausnahme der Hybridisierungen DCM_175_RA, AVkompl_210_RA, HLHS_11 und TA_629 (Abbildung 35). Letztgenannte Proben zeigten einen r-Wert von $r < 0,97$ (Grenzwert $r \geq 0,97$) und wurden daher aus der finalen statistischen Analyse ausgeschlossen, um die Heterogenität innerhalb der Gruppen für den nachfolgenden T-Test zu minimieren. Die Qualitätsfilterung führte somit zum Ausschluss von vier der 68 Hybridisierungen, sodass Rohdaten von insgesamt 64 Herzgewebeproben normalisiert und GS-gelutert wurden. Hierbei wurden 36763 CpG-Loci exkludiert und 448814 CpG-Loci gingen in die finale statistische Analyse ein (Tabelle 27, P3). Von den so prozessierten AVG-Beta-Werten der Replikate der LV-, RV-, LA-, RA-Proben von LV-An und HLHS-TGA wurden dann Mittelwerte gebildet, sodass insgesamt 49 Hybridisierungen, bestehend aus 44 Atrium-Proben (35 IAS, 3 LA, 6 RA) und 5 Ventrikel-Proben (3 LV, 2 RV), in die finale statistische Analyse mittels QluCore-Software gingen.

3.2.2.3.2 Identifizierung von 168 signifikant differenziell methylierten CpG-Loci in Atriumgewebe verglichen mit Ventrikelgewebe

Um zu untersuchen, ob sich die einzelnen Herzgewebesubtypen (IAS, LA, RA, LV und RV) in übergeordnete Gewebegruppen aufteilen ließen, wurde eine **ANOVA** nach Herzgewebetyp durchgeführt. Die ANOVA ergab bei $q \leq 1 \times 10^{-6}$ mit 271 DM-CpG-Loci eine deutliche Trennung der Gewebegruppen „Atrium“ und „Ventrikel“ (Abbildung 36). Die atrialen Gewebetypen IAS, LA und RA und die ventrikulären Gewebetypen LV und RV bildeten jeweils eine Gruppierung. In einem anschließenden T-Test sollte dann die DNA-Methylierung zwischen Atrium- und Ventrikelproben verglichen werden.

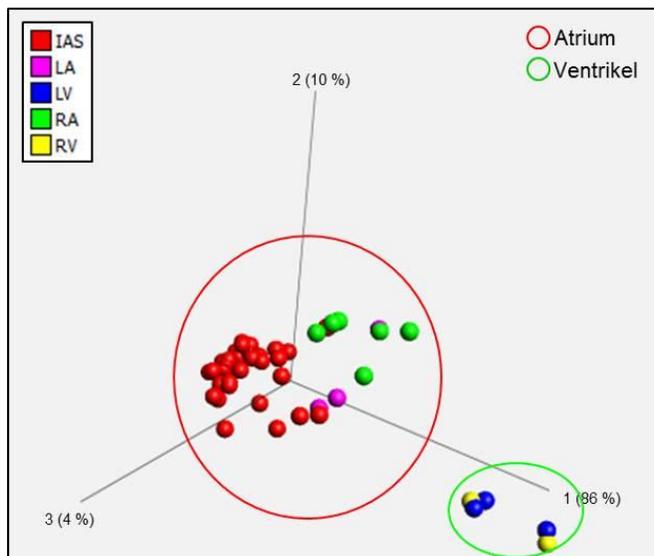


Abbildung 36 PCA-Darstellung einer ANOVA der AVG-Beta-Werte von atrialen (LA, RA, IAS) und ventrikulären (LV, RV) Herzgewebeproben, die mittels 450K-Array analysiert wurden. Dargestellt ist die Segregation der verschiedenen Herzgewebeproben ($N = 49$) basierend auf einer ANOVA: 271/448814 CpG-Loci mit $\sigma/\sigma_{\max} \geq 0,4$ und $q \leq 1 \times 10^{-6}$. Farbcode nach Herzgewebetyp: rot: IAS, rosa: LA, blau: LV, grün: RA, gelb: RV. Die ANOVA nach Herzgewebetyp zeigte eine Segregation von zwei verschiedenen Gewebegruppen „Atrium“ (rote Umrandung) und „Ventrikel“ (grüne Umrandung). Normalisierung zur graphischen Darstellung: $m = 0$.

Der **T-Test** mit $q \leq 1 \times 10^{-6}$ ($\sigma/\sigma_{\max} \geq 0,4$) resultierte in 168 signifikant differenziell methylierten CpG-Loci zwischen „Atrium“ und „Ventrikel“ mit Methylierungsdifferenzen von $\Delta\beta > 30\%$ (Heatmap Abbildung 37). Die atrialen Herzgewebeproben zeigten in der Heatmap bei ca. 75 % der DM-CpG-Loci eine Hypomethylierung im Vergleich zu ventrikulären Proben, während die restlichen 25 % der DM-CpG-Loci in Atrium-Proben verglichen mit Ventrikel-Proben tendenziell hypermethyliert waren. Insgesamt lag eine signifikante Anreicherung von RDMR-assoziierten CpG-Loci innerhalb der 168 signifikant differenziell methylierten CpG-Loci zwischen „Atrium“ und „Ventrikel“ vor ($OR = 4,24$, $p = 2,1 \times 10^{-6}$). Auch Enhancer-assoziierte CpG-Loci zeigten mit $OR = 3,63$ ($p = 5,9 \times 10^{-16}$) eine signifikante Anreicherung.

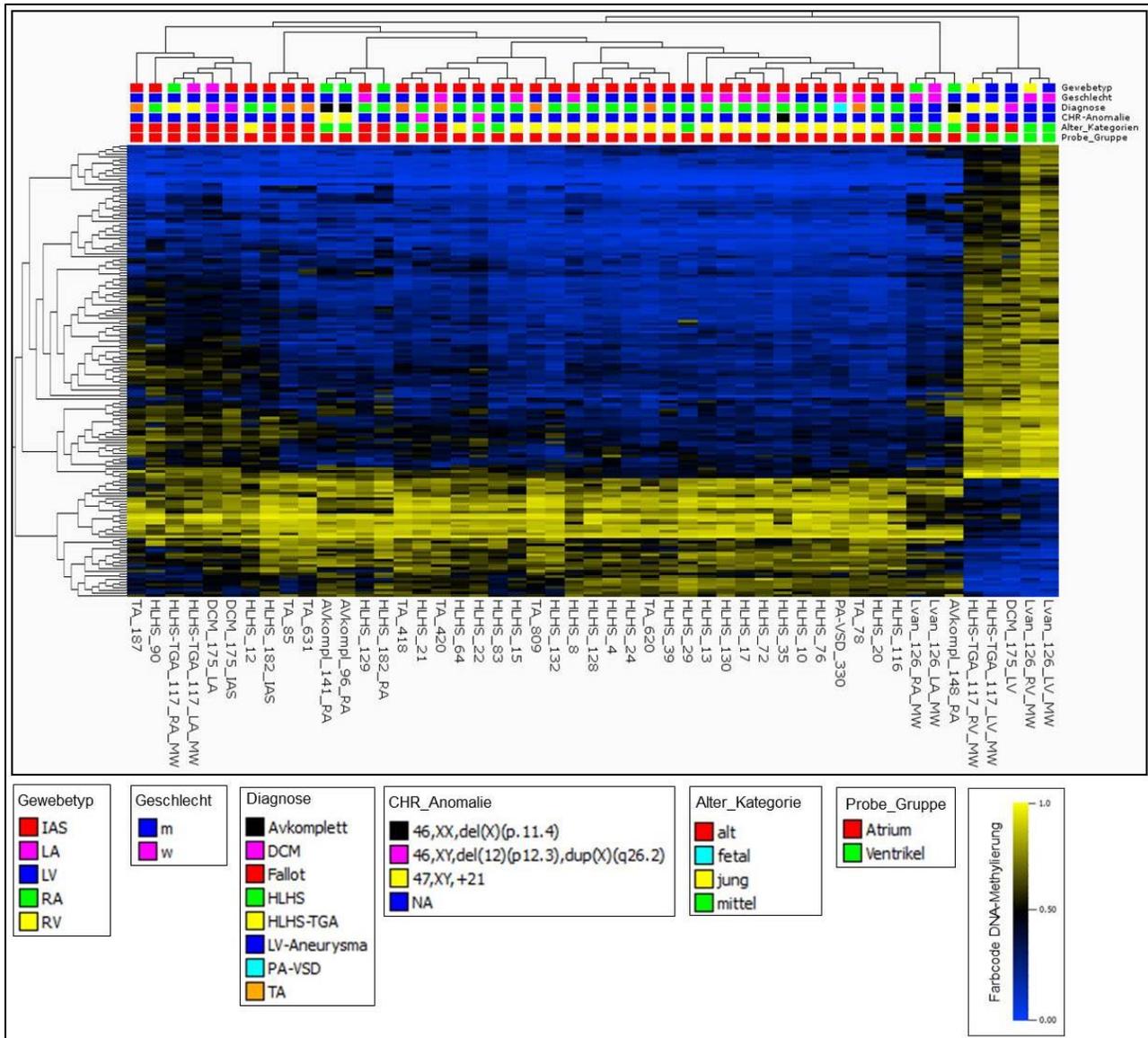


Abbildung 37 Heatmap der AVG-Beta-Werte des T-Tests der Gruppen „Atrium“ (LA, RA, IAS) und „Ventrikel“ (LV, RV) bestehend aus 49 Herzgewebeproben von Patienten mit CHDs, die mittels 450K-Array analysiert wurden. AVG-Beta-Mittelwerte sind hierbei von den Proben Lvan und HLHS-TGA eingegangen (siehe Beschriftung „MW“). Dargestellt ist die Heatmap des T-Tests der Gruppen „Atrium“ vs. „Ventrikel“ mit $\sigma/\sigma_{max} \geq 0,4$ und $q \leq 1 \times 10^{-6}$, wobei 168/448814 CpG-Loci differenziell methyliert in „Atrium“ im Vergleich zu „Ventrikel“ vorlagen. Jede Zeile stellt einen CpG-Locus dar, mit niedrigen AVG-Beta-Werten in blau und hohen AVG-Beta-Werten in gelb. Jede Spalte stellt eine DNA-Probe dar, mit den Bezeichnungen der einzelnen Proben unterhalb der Heatmap. Proben und CpG-Loci wurden hierarchisch geclustert. Oberhalb der Heatmap sind pro Probe entsprechend der Farbcode-Legende unterhalb der Heatmap folgende Kategorien farblich dargestellt (von oben nach unten): der Gewebetyp, das Geschlecht, die kardiale Diagnose, bekannte chromosomale Anomalien (CHR-Anomalie), Alterskategorien („jung“ (≤ 3 Wochen), „mittel“ (> 3 Wochen bis ≤ 30 Wochen) und „alt“ (ab > 30 Wochen)) und die Probengruppenkategorie. Normalisierung zur graphischen Darstellung: keine, Einfärbung nach absoluten AVG-Beta-Werten.

Von den 168 zwischen den beiden Gruppen „Atrium“ und „Ventrikel“ signifikant differenziell methylierten CpG-Loci konnten 101 CpG-Loci insgesamt 78 Genen zugeordnet werden. CHD-assoziierte Gene, die multiple DM-CpGs (3 DM-CpG-Loci) aufwiesen, waren *IRX4* und *FGFR2*.

Mutationen in *IRX4* konnten zuvor bei Patienten mit CHDs gefunden werden [332]. Das Gen *FGFR2* wurde in Zusammenhang mit dem Apert-Syndrom, welches in 10 % der Fälle in Assoziation mit CHDs auftrat [322], beschrieben. Im *Welch Two Sample* T-Test (wT-Test im Folgenden abgekürzt) lag bei allen DM-CpG-Loci von *IRX4* (cg17774559: $p = 1,7 \times 10^{-5}$, cg04180086: $p = 0,0001$; cg18172877: $p = 4,5 \times 10^{-5}$) und *FGFR2* (cg17710779: $p = 0,0005$, cg01385327: $p = 0,0007$; cg10362842: $p = 1,5 \times 10^{-13}$) eine hoch signifikante differenzielle Methylierung zwischen den Gruppen „Atrium“ und „Ventrikel“ vor (Abbildung 38). Die atrialen Proben wiesen hierbei eine Hypomethylierung im Vergleich zu den ventrikulären Proben auf.

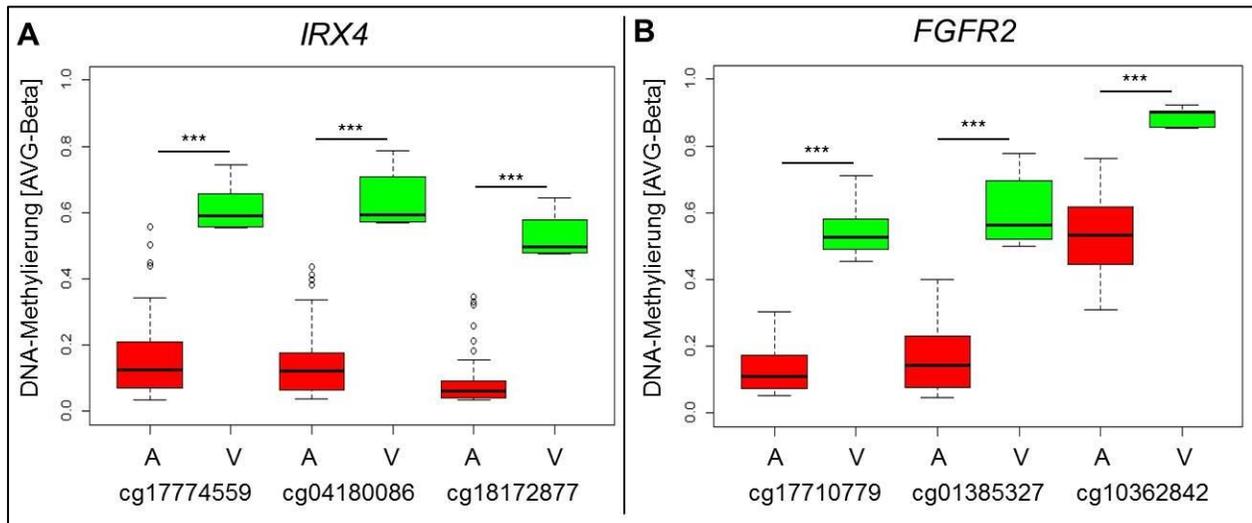


Abbildung 38 Boxplot-Darstellung der je drei DM-CpG-Loci der Gene (A) *IRX4* und (B) *FGFR2*, welche in Atriumgewebe („A“) differenziell methyliert im Vergleich zu Ventrikelgewebe („V“) vorlagen. Die Gruppe „Atrium“ umfasste 44 Herz-Gewebeproben und die Gruppe „Ventrikel“ umfasste fünf Herzgewebeproben. In den Boxplotdarstellungen A und B sind pro Gen und CpG-Locus die DNA-Methylierung (AVG-Beta-Werte) auf der y-Achse und die Probenbezeichnung auf der x-Achse dargestellt. Boxplots der Atrium-Gruppe sind rot und die der Ventrikel-Gruppe grün eingefärbt. Die p-Werte des wT-Tests sind jeweils pro analysiertem CpG nach Sternchencode⁴² aufgeführt. Alle DM-CpG-Loci zeigten eine hoch signifikante ($p < 0,001$) differenzielle Methylierung zwischen den Gruppen „Atrium“ und „Ventrikel“ (Grenzwert: $p \leq 0,05$ (*)).

Innerhalb der 78 Gene mit DM-CpGs war nach GO-Analyse (siehe 2.2.5.3) eine signifikante **Anreicherung** von Genen im biologischen Prozess „cardiovascular system development (GO:0072358)“ (13 Gene, $q = 6 \times 10^{-4}$) festzustellen. Der GO-Prozess beinhaltete auch die Gene *IRX4* und *FGFR2*, die multiple DM-CpGs zwischen atrialen und ventrikulären Herzgewebeproben aufzeigten.

⁴² Sternchencode p-Werte: $p < 0,05$ entspricht * (knapp signifikant), $p < 0,01$ entspricht ** (signifikant), $p < 0,001$ entspricht *** (hoch signifikant)

3.2.2.3.3 Verifizierung

Zur Verifizierung der differentiellen DNA-Methylierung zwischen atrialen und ventrikulären Herzgewebetypen wurde je ein CpG-Locus der Gene *IRX4* (cg17774559) und *FGFR2* (cg17710779) mittels Epityper-Technik (siehe 2.2.2.4) analysiert (Primer siehe 2.1.2.1 Tabelle 4). Die Qualitätskontrolle der mittels *in silico* Assay-Vorhersage ermittelten Konversionskontrollen (Details siehe 2.2.5.1.1) ergab keine *Peak*-Detektion bei den analysierten Cytosinen im nicht-CpG-Kontext, sodass von einer effizienten Bisulfit-Konversion ausgegangen werden konnte. Hierbei wurden für das „Atrium“-Kollektiv (Summe = 6 Proben) pro HLHS-TGA- und LV-An-Patient jeweils die Probe Nr. 1 der LA/RA-Replikate, sowie die IAS-Proben HLHS_4_IAS und HLHS_182_IAS eingesetzt. Das „Ventrikel“-Kollektiv (Summe = 4 Proben) setzte sich jeweils aus der Probe Nr. 1 der LV/RV-Replikate von HLHS-TGA- und LV-An-Patient zusammen. Zur Analyse der Korrelation der Techniken 450K-Array und Epityper wurden die Methylierungswerte der Epityper-Analyse auf der y-Achse gegen die Methylierungswerte des 450K-Arrays auf der x-Achse in einem Scatterplot als lineare Regression (Abbildung 39) dargestellt. Insgesamt fielen hierbei die Methylierungswerte der Epityper-Technik geringer aus als die 450K-Methylierungswerte, mit jedoch vergleichbaren Methylierungsdifferenzen (Rohdaten siehe Anhang (Abschnitt 8, Abbildung 61)). Das Bestimmtheitsmaß R^2 ergab für *IRX4* cg17774559 $R^2 = 0,94$ und für *FGFR2* cg17710779 $R^2 = 0,85$. *IRX4* zeigte eine hohe Korrelation mit R^2 nahe 1, während der analysierte CpG-Locus von *FGFR2* eine etwas geringere Korrelation der Daten ($R^2 = 0,85$) aufwies. Letzterer lag jedoch nahe dem Grenzwert $R^2 \geq 0,9$, sodass die Verifizierung als valide eingestuft wurde. Die differentielle DNA-Methylierung zwischen atrialen und ventrikulären Herzgewebeproben konnte somit verifiziert werden.

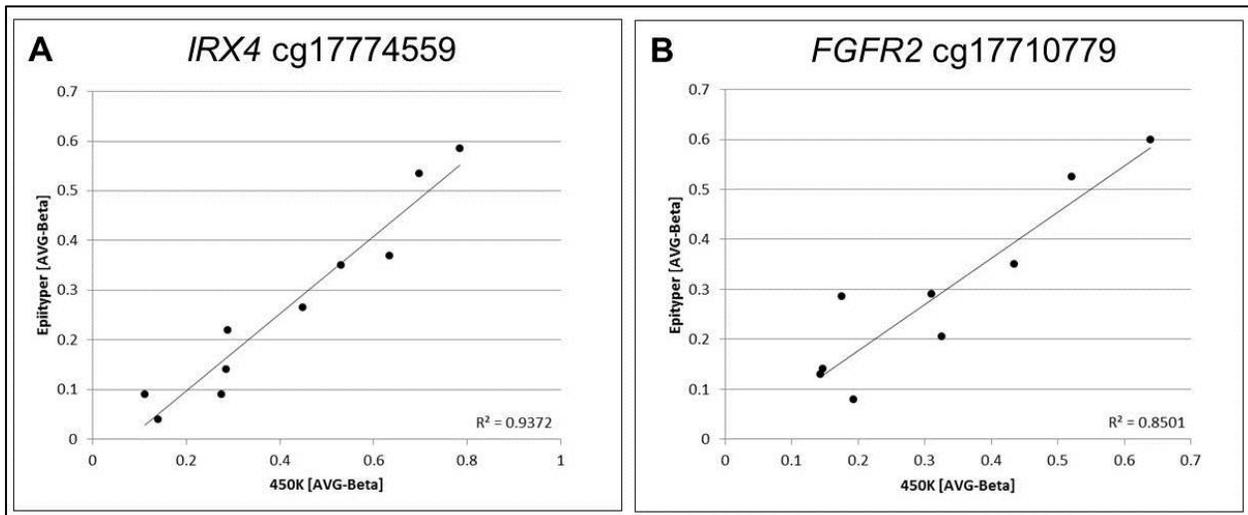


Abbildung 39 Lineare Regression der DM-CpGs der Gene (A) *IRX4* (cg17774559) und (B) *FGFR2* (cg17710779) der Korrelationsanalyse (Pearson-Methode) der DNA-Methylierungswerte des 450K-Arrays und der Epityper-Analyse. Die aufgeführten CpG-Loci lagen signifikant differenziell methyliert zwischen atrialen und ventrikulären Herzgewebetypen in der 450K-Analyse vor. Zur Verifizierung mittels Epityper-Technik wurden 6 DNA-Proben des „Atrium“-Kollektivs (pro HLHS-TGA- und LV-An-Patient jeweils die Probe Nr. 1 der LA/RA-Replikate, sowie die IAS-Proben HLHS_4_IAS und HLHS_182_IAS) und vier DNA-Proben des „Ventrikel“-Kollektivs (Probe Nr. 1 der LV/RV-Replikate von HLHS-TGA- und LV-An-Patient) analysiert. Methylierungswerte der Epityper-Analyse auf der y-Achse wurden gegen die 450K-Methylierungswerte auf der x-Achse in einem Scatterplot als lineare Regression dargestellt. Das Bestimmtheitsmaß R^2 ergab für cg17774559 $R^2 = 0,94$ und für cg17710779 $R^2 = 0,85$.

Zusammenfassend konnte bei 168 CpG-Loci eine signifikante differenzielle DNA-Methylierung mit einer Hypomethylierung von 75 % der DM-CpG-Loci im Atriumgewebe im Vergleich zum Ventrikelgewebe detektiert werden. Eine Unterscheidung dieser Gewebegruppen ist somit anhand der detektierten 168 CpG-Loci möglich. Es lag eine Anreicherung von RDMR- und Enhancer-assoziierten CpG-Loci vor, sodass vornehmlich regulative DNA-Regionen in Atriumgewebe verglichen mit Ventrikelgewebe differenziell methyliert waren. Eine differenzielle Methylierung multipler CpG-Loci mit Methylierungsdifferenzen von $\Delta\beta > 30\%$ Herzentwicklungsassoziierter Gene konnte in den Genen *IRX4* und *FGFR2* mit einer signifikanten Hypomethylierung in Atriumgewebe verglichen mit Ventrikelgewebe detektiert und verifiziert werden.

3.2.2.4 Herzmuskelgewebe und Gefäßstrukturen des Herzens zeigen distinkte DNA-Methylierungs-Signaturen bei einem Patienten mit DCM

Beim vorausgegangenen Herz-Mapping des Teilprojekts P2 (siehe 3.2.2.2) konnte bereits in der unsupervidierten Analyse eine separate Gruppierung von Gefäß-/„gefäß-ähnlichen“ Gewebeproben detektiert werden (Abbildung 32). In diesem Teilprojekt P4 sollte daher untersucht werden, ob signifikante Unterschiede in der DNA-Methylierung zwischen Gefäß- und muskulärem Gewebe des Herzens nachweisbar sind. Zudem sollte die Fragestellung beantwortet werden, welche *Gen-Pathways* mit einer differenziellen Methylierung angereichert waren.

In diese Analyse gingen ausschließlich die neun Herzgewebeproben des DCM-Patienten (IAS, LA, RA, LV, MV, TV, PA, Aorta und AoV (siehe 2.1.1.1, Tabelle 1)) ein. So konnten DNA-Methylierungseffekte durch unterschiedliche Individuen, sowie etwaige Effekte durch unterschiedliche Gewebeisolationen (*post mortem* vs. frisch isoliertes Gewebe) ausgeschlossen werden. Zudem sind diese Proben auf demselben Array prozessiert worden, sodass auch Batch-Effekte ausgeschlossen werden konnten.

3.2.2.4.1 Qualitätskontrolle

Um zu analysieren, ob eine Unterscheidung verschiedener Gewebegruppen bereits mithilfe der Erhebung von Korrelationskoeffizienten möglich war, wurde eine CpG-Gesamtkorrelation durchgeführt.

Die CpG-Gesamtkorrelation ergab bei Proben von Gefäßen (PA, Aorta) bzw. „gefäßähnlichem“ Gewebe (AoV) deutlich abweichende *r*-Werte im Vergleich zu den restlichen Herzgewebeproben muskulärem Ursprungs des DCM-Patienten (Abbildung 40). Innerhalb der jeweiligen Gewebegruppen „Gefäß“ und „Muskel“ des Herzens lagen hierbei homogen hohe *r*-Werte (nahe $r = 1$) vor.

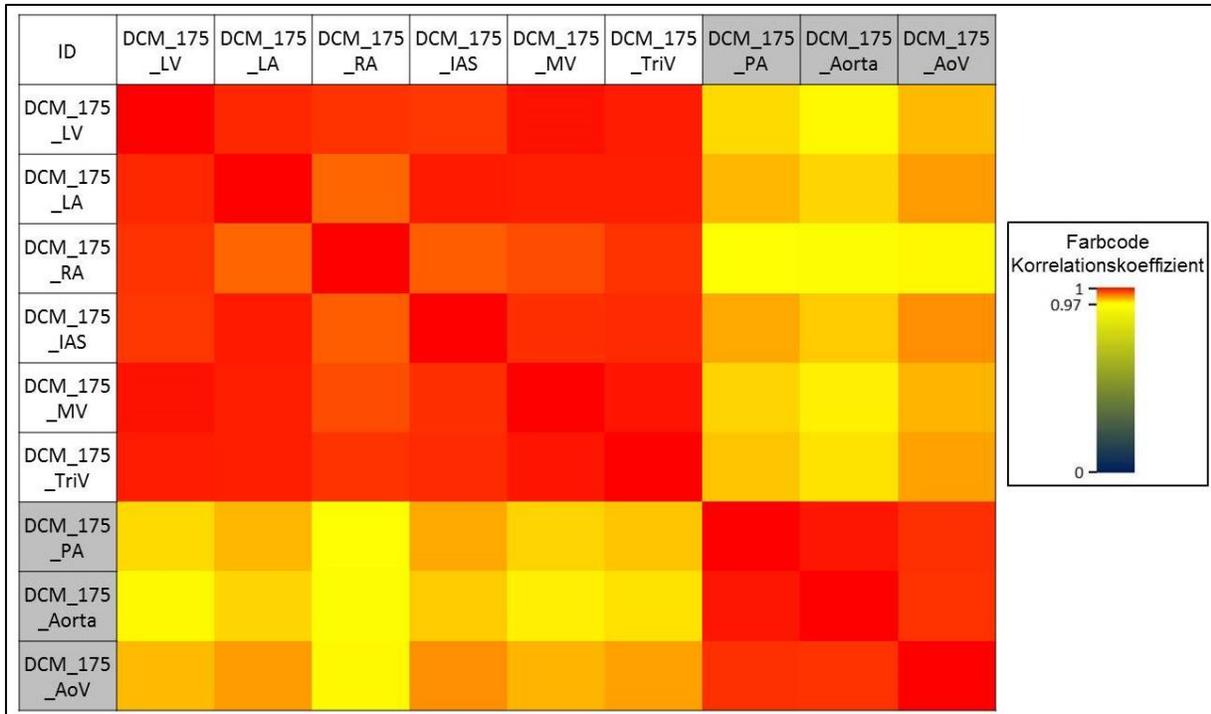


Abbildung 40 CpG-Gesamtkorrelation der AVG-Beta-Werte von neun Herzgewebeprobe eines DCM-Patienten (IAS, LA, RA, LV, MV, TriV, PA, Aorta und AoV), die auf einem 450K-Array hybridisiert wurden. Dargestellt sind die Korrelationskoeffizient-Werte (r) nach Pearson-Methode mit einer farblich bedingten Formatierung von niedrigen r-Werten in blau und hohen r-Werten in rot (siehe Farbcode). Jede Spalte (und die gleiche Zeilennummer) steht für eine DNA-Probe, mit oberhalb und links der Korrelationsmatrix den Herzgewebeprobe aus „Muskel“ in weiß und „Gefäß“-Probe in grau.

Die Normalisierung und GS-Filterung der 450K-Rohdaten der neun DCM-Proben resultierte in einem Ausschluss von 35328 CpG-Loci und 450249 CpG-Loci gingen in die finale statistische Analyse ein (Tabelle 27, P4).

3.2.2.4.2 Identifizierung von 772 signifikant differenziell methylierten CpG-Loci in Herzmuskelgewebe im Vergleich zu Gefäßstrukturen des Herzens

Wie zuvor schon in der CpG-Gesamtkorrelation erkennbar, ergab auch die **unsupervidierte Analyse** der 450249 CpG-Loci (Varianzfilter $\sigma/\sigma_{\max} \geq 0,4$) der neun DCM-Proben eine deutliche Gruppierung von „Muskel“- und „Gefäß“-Proben in der PCA-Darstellung (Abbildung 41). Die AoV-Gewebeprobe, physiologisch den Herzklappen zuzuordnen, wurde hier als „Gefäß“ definiert (Abbildung 41 grüne Umrandung), da diese Probe bei der Gewebeexzision aus dem Aortenklappenring entnommen wurde (siehe 2.1.1.1). Die Herzklappen MV und TV wurden unter

der Gruppe „Muskel“ zusammengefasst, da die beiden Proben mit den muskulären Gewebeproben LA, RA und LV gruppierten (Abbildung 41 rote Umrandung).

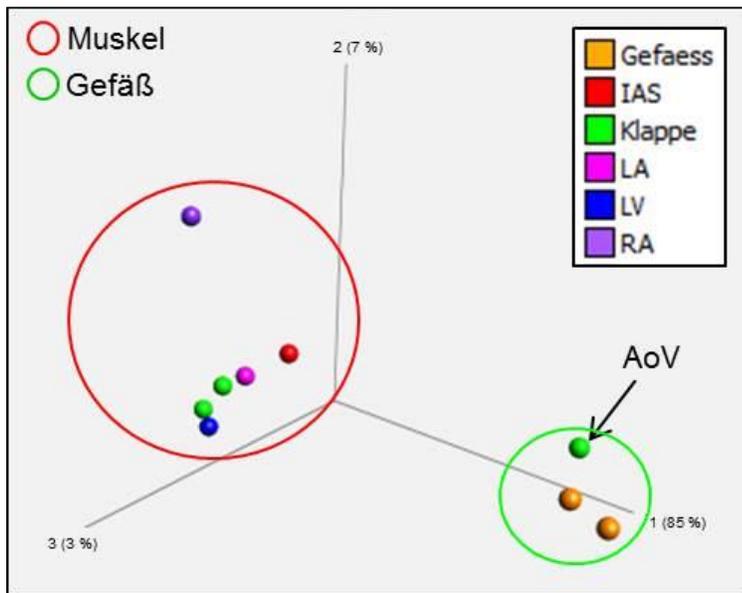


Abbildung 41 PCA-Darstellung der unsupervidierten Analyse von neun Herzgewebeproben eines DCM-Patienten (IAS, LA, RA, LV, MV, TV, PA, Aorta und AoV), die mittels 450K-Array analysiert wurden. Dargestellt ist die Segregation der verschiedenen Herzgewebeproben basierend auf 12363/450249 CpG-Loci mit $\sigma/\sigma_{\max} \geq 0,4$. Farbcode Herzgewebetyp: orange: Gefäßproben (PA, Aorta), rot: IAS, grün: Herzklappen (MV, TV, AoV) rosa: LA, blau: LV, violett: RA. In der unsupervidierten Analyse war eine deutliche Gruppierung von „Muskel“-Herzgewebeproben (rote Umrandung) und Gefäß-/„gefäß-ähnlichen“-Proben“ (grüne Umrandung) zu erkennen. Normalisierung zur graphischen Darstellung: $m = 0$

Ein **T-Test** mit $q \leq 1 \times 10^{-5}$ ($\sigma/\sigma_{\max} \geq 0,4$) resultierte in 772 CpG-Loci mit Methylierungsdifferenzen von $\Delta\beta > 30\%$ zwischen „Muskel“ und „Gefäß“ (Heatmap Abbildung 42). 660 CpG-Loci und damit 85 % der DM-CpG-Loci zeigten eine Hypermethylierung in „Muskel“ im Vergleich zu „Gefäß“. Insgesamt lag eine signifikante Anreicherung von Enhancer-assoziierten CpG-Loci innerhalb der 772 DM-CpG-Loci zwischen „Muskel“ und „Gefäß“ vor (OR = 3,6 $p = 2,2 \times 10^{-16}$). Auch RDMR-assoziierte CpG-Loci zeigten mit OR = 2,22 ($p = 4,1 \times 10^{-6}$) eine signifikante Anreicherung innerhalb der DM-CpG-Loci in Herzmuskelgewebe verglichen mit Gefäßstrukturen.

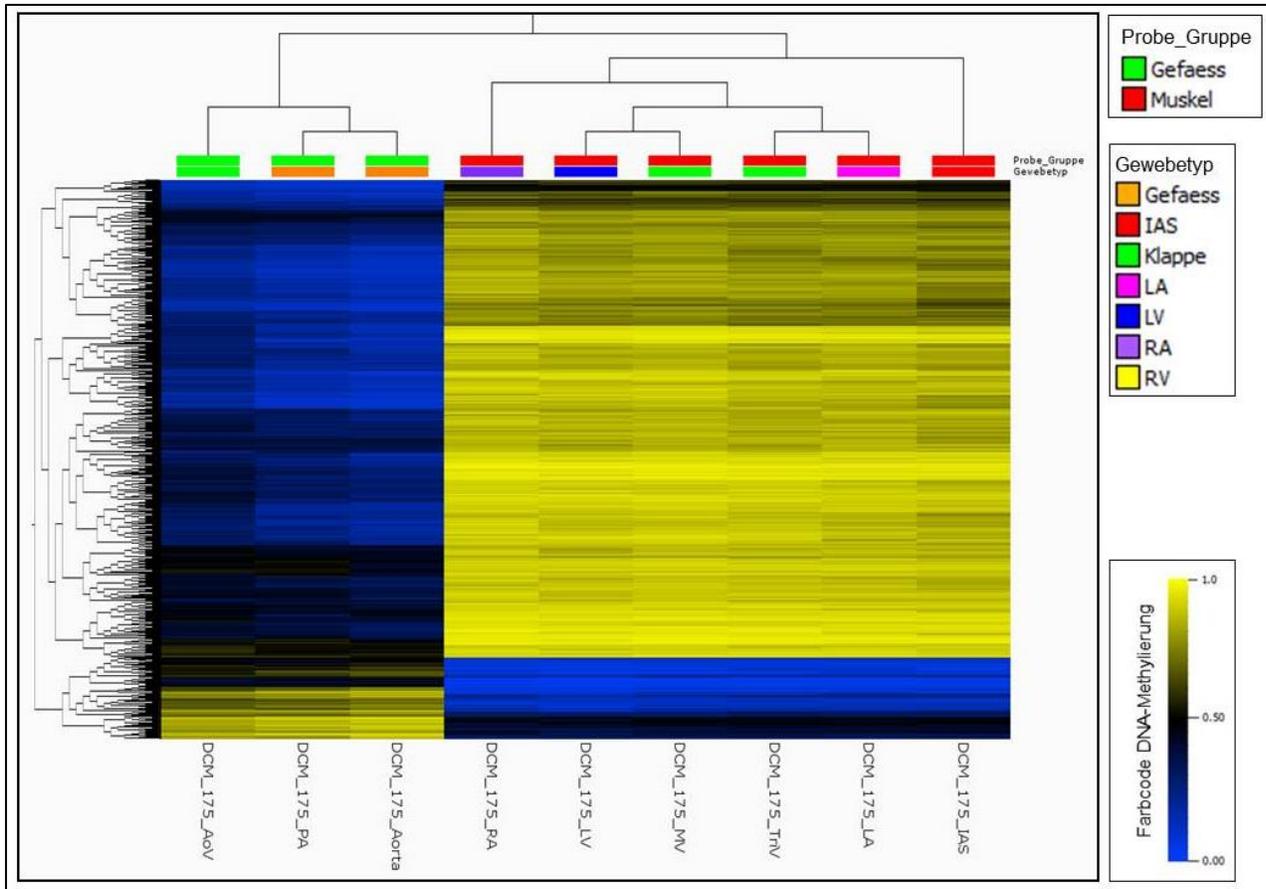


Abbildung 42 Heatmap der AVG-Beta-Werte des T-Tests der Gruppen „Muskel“ (LA, RA, IAS, LV, RV, TV, MV) und „Gefäß“ (PA, Aorta, AoV) bestehend aus Herzgewebeprobe von einem DCM-Patienten, die mittels 450K-Array analysiert wurden. Dargestellt ist die Heatmap des T-Tests der Gruppen „Muskel“ vs. „Gefäß“ mit $\sigma/\sigma_{max} \geq 0,4$ und $q \leq 1 \times 10^{-5}$, wobei 772/450249 CpG-Loci differenziell methyliert in „Muskel“ im Vergleich zu „Gefäß“ vorlagen. Jede Zeile stellt einen CpG-Locus dar, mit niedrigen AVG-Beta-Werten in blau und hohen AVG-Beta-Werten in gelb. Jede Spalte stellt eine DNA-Probe dar, mit den Bezeichnungen der einzelnen Proben unterhalb der Heatmap. Proben und wurden hierarchisch geclustert. Oberhalb der Heatmap sind pro Probe entsprechend der Farbcode-Legenden rechts der Heatmap folgende Kategorien farblich dargestellt (von oben nach unten): die Probengruppenkategorie und der Gewebetyp (Gefäß: PA, Aorta; Herzklappen: MV, TriV, AoV). Normalisierung zur graphischen Darstellung: keine, Einfärbung nach absoluten AVG-Beta-Werten.

Von den 772 signifikant differenziell methylierten CpG-Loci zwischen „Muskel“ und „Gefäß“ konnten 548 CpG-Loci insgesamt 421 Genen zugeordnet werden. Die in „Muskel“ hypermethylierten 660 CpG-Loci im Vergleich zu „Gefäß“ zeigten innerhalb der 369 assoziierten Gene nach GO-Analyse (siehe 2.2.5.3) eine signifikante **Anreicherung** von Genen mit der molekularen Funktion „actin binding (GO:0003779)“ (26 Gene, $q = 1,25 \times 10^{-5}$). Die 55 Gene mit in „Muskel“ hypomethylierten 112 CpG-Loci im Vergleich zu „Gefäß“ ergaben keine signifikante Anreicherung.

Die Gene *GRIK2* (11 CpGs), *DIP2C* (10 CpGs) und *PRDM16* (9 CpGs) zeigten mit 9 bis 11 CpG-Loci eine besonders hohe Anzahl von DM-CpG-Loci pro Gen (weitere Gene

umfassten max. 4 DM CpG-Loci). Die DM-CpG-Loci in *GRIK2* lagen im Herzmuskelgewebe verglichen mit Gefäßstrukturen mit mittleren Methylierungsdifferenzen von $\Delta\beta = 47\%$ signifikant hypomethyliert im wT-Test vor ($p = 2,5 \times 10^{-9}$). Die DM-CpG-Loci von *DIP2C* ($p = 9,5 \times 10^{-10}$) und *PRDM16* ($p = 5,6 \times 10^{-8}$) zeigten jeweils mittlere Methylierungsdifferenzen von $\Delta\beta = 61\%$ bzw. 53% mit einer Hypermethylierung im Herzmuskelgewebe im Vergleich zu Gefäßstrukturen (Abbildung 43).

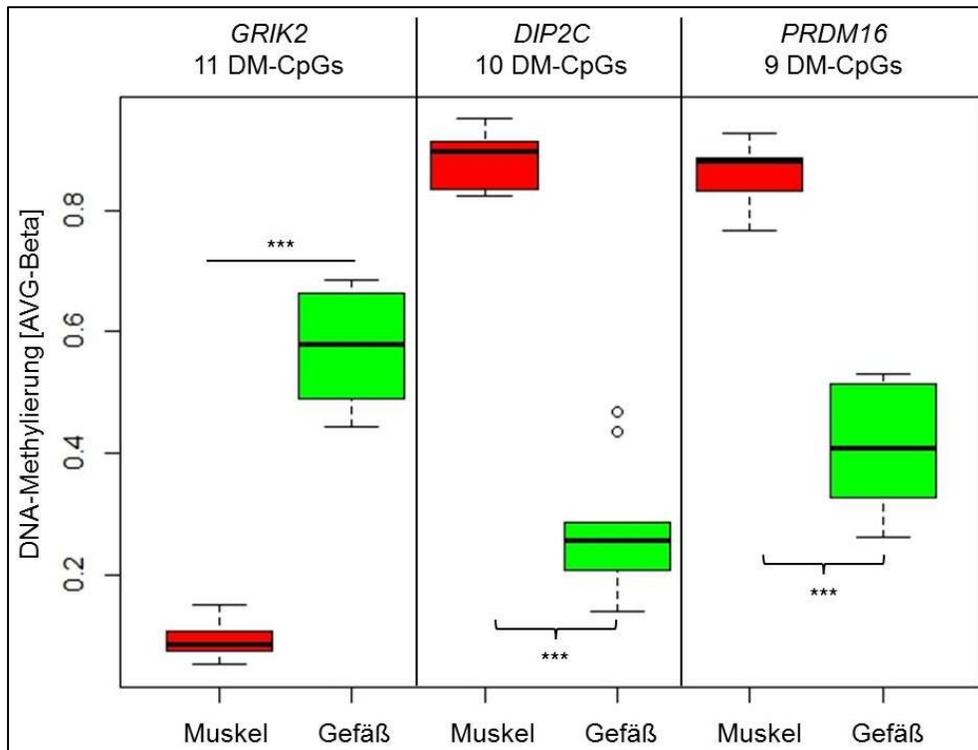


Abbildung 43 Boxplot-Darstellung multipler DM-CpG-Loci der Gene *GRIK2*, *DIP2C* und *PRDM16*, welche in „Muskel“ differenziell methyliert im Vergleich zu „Gefäß“ vorlagen. Die Gruppe „Muskel“ umfasste sieben Herz-Gewebeproben (LA, RA, IAS, LV, RV, TV, MV) und die Gruppe „Gefäß“ umfasste drei Herzgewebeproben (PA, Aorta, AoV). In der Boxplotdarstellung sind pro Gen die mittleren DNA-Methylierungswerte (AVG-Beta-Werte) der DM-CpG-Loci pro Gruppe (Muskel/Gefäß) auf der y-Achse und die Probenbezeichnung auf der x-Achse dargestellt. Boxplots der Muskel-Gruppe sind rot und die der Gefäß-Gruppe grün eingefärbt. Die p-Werte des wT-Tests sind jeweils pro analysiertem CpG nach Sternchencode aufgeführt. Alle DM-CpG-Loci zeigten eine hoch signifikante ($p < 0,001$) differenzielle Methylierung zwischen den Gruppen „Muskel“ und „Gefäß“ (Grenzwert: $p \leq 0,05$ (*)).

Die Analyse der DNA-Methylierungs-Signatur von muskulärem Gewebe des Herzens im Vergleich zu Gefäßstrukturen ergab zusammenfassend 772 signifikant differenziell methylierte CpG-Loci mit Methylierungsdifferenzen von $\Delta\beta > 30\%$. Die detektierten 772 DM-CpG-Loci ermöglichen somit eine Unterscheidung von muskulärem Gewebe und Gefäßstrukturen des Herzens anhand der DNA-Methylierungsmuster. Interessanterweise lag der Großteil (85 %) der DM-CpG-Loci in „Muskel“ hypermethyliert im Vergleich zu „Gefäß“ vor. Mit einer signifikanten Anreicherung von RDMR- und Enhancer-assoziierten CpG-Loci zeigten vornehmlich regulative DNA-Regionen eine differenzielle DNA-Methylierung zwischen „Muskel“ und „Gefäß“.

3.2.2.5 Identifizierung von 86 differenziell methylierten CpG-Loci in Herzgewebe von HLHS-Patienten im Vergleich zu fetalem Herzkontrollgewebe

Um das DNA-Methylierungsprofil von IAS-Gewebe von HLHS-Patienten mit gesunden Kontrollindividuen vergleichen zu können, wurde als Kontrollkollektiv fetales Herzgewebe gesunder Individuen (Biocat, (Heidelberg, Deutschland)) von je einem weiblichen (40. Woche) und einem männlichen Donor (25. Woche) verwendet. So sollten krankheitsspezifische Effekte, die möglicherweise in der initialen Analyse des Vergleichs der Gruppen HLHS und Rechtsherzfehlbildungen (siehe 3.2.2.1) von Einfluss waren, vermieden werden. Um zudem altersspezifische DNA-Methylierungseffekte zu minimieren, wurden in dieser Analyse des Teilprojekts P5 lediglich die IAS-Gewebeproben der Alterskategorie „jung“ (≤ 3 Wochen) des HLHS-Probenkollektivs eingesetzt, um diese mit dem fetalen Herzkontrollgewebe zu vergleichen.

In diese Analyse gingen somit 17 IAS-Gewebeproben von HLHS-Patienten der Alterskategorie „jung“ (≤ 3 Wochen, siehe 2.1.1.1 Tabelle 1 „J“) und Duplikate von fetalem Herzgewebe eines männlichen und weiblichen Donors als Kontrollen ein (Tabelle 1, P5).

3.2.2.5.1 Qualitätskontrolle

Die CpG-Gesamtkorrelation der AVG-Beta-Werte der insgesamt 21 Hybridisierungen ergab hohe r-Werte mit einem Median über alle Korrelationskoeffizienten von $r \geq 0,98$, sodass keine Proben exkludiert werden mussten (siehe Abschnitt 8 Abbildung 62). Die Duplikate der fetalen Herzgewebeproben wiesen wie erwartet einen hohen Korrelationskoeffizienten von $r > 0,99$ zueinander auf. Nach Normalisierung und GS-Filterung der 450K-Rohdaten der insgesamt 21 Hybridisierungen wurden 35543 CpG-Loci ausgeschlossen und 450034 CpG-Loci gingen in die finale statistische Analyse ein (Tabelle 27, P4). Von den so prozessierten AVG-Beta-Werten der Duplikate der fetalen Herzgewebeproben wurden Mittelwerte gebildet, sodass insgesamt 19 Hybridisierungen, bestehend aus 17 IAS-Gewebeproben von Patienten mit HLHS und zwei fetalen Herzgewebeproben gesunder Individuen final analysiert wurden.

3.2.2.5.2 Identifizierung multipler differenziell methylierter CpG-Loci der Gene

NR2F1, *NR2F1-AS1*, *PDXDC1* und *XYLT1* in Herzgewebe von HLHS-Patienten im Vergleich zu fetalem Herzkontrollgewebe

Es ließen sich 137 signifikant differenziell methylierte CpG-Loci bei $q \leq 1 \times 10^{-6}$ im T-Test (basierend auf $\sigma/\sigma_{\max} \geq 0,4$) im Herzgewebe von HLHS-Patienten im Vergleich zu fetalem Herzkontrollgewebe detektieren (Abbildung 44). Bei ca. 50 % der DM-CpG-Loci konnte eine Hypermethylierung im IAS-Gewebe von Patienten mit HLHS verglichen mit fetalem Herzkontrollgewebe mit hohen DNA-Methylierungsdifferenzen von $\Delta\beta > 45\%$ feststellen.

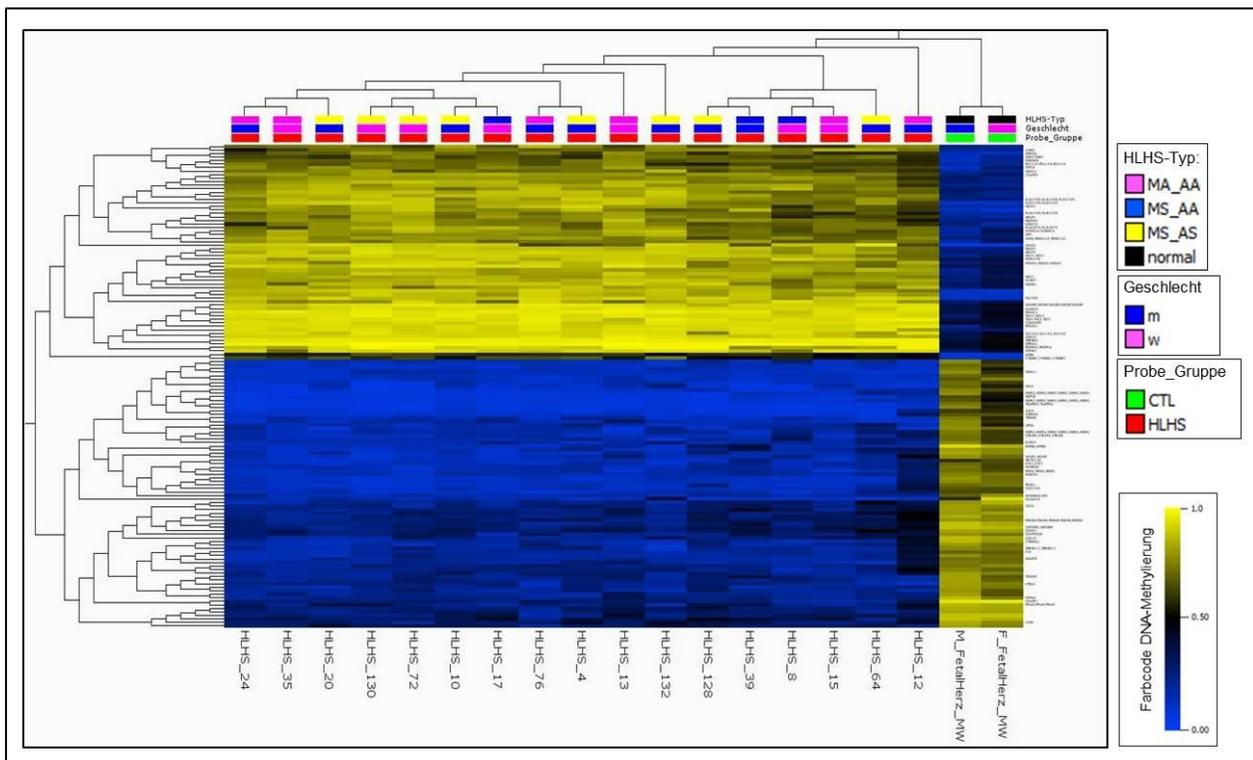


Abbildung 44 Heatmap der AVG-Beta-Werte des T-Tests des HLHS-Kollektivs (17 IAS-Gewebeproben) und des Kontrollkollektivs aus fetalem Herzkontrollgewebe (2 fetale Herzgewebeproben), welche mittels 450K-Array analysiert wurden. Dargestellt ist die Heatmap des T-Tests der Gruppen „HLHS“ vs. „CTL“ (fetale Herzgewebeproben) mit $\sigma/\sigma_{\max} \geq 0,4$ und $q \leq 1 \times 10^{-6}$, wobei 137/450034 CpG-Loci differenziell methyliert bei „HLHS“ im Vergleich zu „CTL“ vorlagen. Jede Zeile stellt einen CpG-Locus dar, mit niedrigen AVG-Beta-Werten in blau und hohen AVG-Beta-Werten in gelb. Jede Spalte stellt eine DNA-Probe dar, mit den Bezeichnungen der einzelnen Proben unterhalb der Heatmap. Proben und CpG-Loci wurden hierarchisch geclustert. Oberhalb der Heatmap sind pro Probe entsprechend der Farbcodierung rechts der Heatmap folgende Kategorien farblich dargestellt (von oben nach unten): der HLHS-Typ („CTL“-Proben als „normal“ bezeichnet), das Geschlecht und die Probenkategorie. Normalisierung zur graphischen Darstellung: keine, Einfärbung nach absoluten AVG-Beta-Werten.

Die fetalen Herzgewebeproben, die nach Angaben des Anbieters (Biocat (Heidelberg, Deutschland)) lediglich mit der Angabe „Gesamtherzgewebe“ beschrieben waren, wurden im Rahmen der Herz-Mapping-Analyse des Teilprojekts P2 (siehe 3.2.2.2) als Herzgewebe

ventrikulärem Ursprungs eingeordnet. Da die Herzgewebeproben der HLHS-Probengruppe aus IAS-Gewebe und damit aus atrialem Herzgewebe bestanden, war das Ziel, „Atrium-Ventrikel“-spezifische DNA-Methylierungseffekte auszuschließen. Demnach wurden im nächsten Schritt CpG-Loci ausgeschlossen, die eine Überlappung zu den zwischen „Atrium“ und „Ventrikel“ differenziell methylierten 168 CpG-Loci (siehe 3.2.2.3) aufwiesen. Dieser Filtervorgang resultierte in 86 „HLHS-spezifischen“ DM-CpG-Loci (Abbildung 45). Auch diese zeigten sehr hohe DNA-Methylierungsdifferenzen von $\Delta\beta > 45\%$ zwischen den Gruppen HLHS und fetalem Herzkontrollgewebe. Diese wurden mit den im Herzgewebe altersabhängig differenziell methylierten 1104 CpG-Loci der HLHS-RF-Analyse aus Abschnitt 3.2.2.1 abgeglichen, um Alterseffekte zu vermeiden. Hierbei ergab sich keine Überschneidung der CpGs. Da die fetalen Herzkontrollgewebe eine gewisse Heterogenität innerhalb der Gruppe aufwiesen und ein Vergleich von lediglich $N = 2$ Herzkontrollgewebe gegen $N = 17$ IAS-Gewebe stattfand, wurden zur Erhöhung der statistischen Power ausschließlich **Gene mit multiplen DM-CpG-Loci** selektiert. Dies resultierte in folgenden vier Genen mit je zwei DM-CpG-Loci (Abbildung 45): *NR2F1*, *NR2F1-AS1*, *PDXDC1* und *XYLT1*.

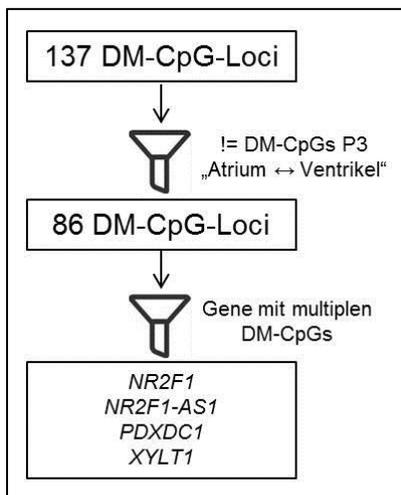


Abbildung 45 Filter-Schema der 137/450034 DM-CpG-Loci im Herzgewebe von HLHS-Patienten im Vergleich zu fetalem Herzkontrollgewebe. Nach Filterung gegen DM-CpG-Loci aus Teilprojekt P3 (siehe 3.2.2.3), die zwischen „Atrium“ und „Ventrikel“ differenziell methyliert vorlagen, resultierten 86 „HLHS-spezifische“ DM-CpG-Loci. Vier Gene (*NR2F1*, *NR2F1-AS1*, *PDXDC1* und *XYLT1*) zeigten hiervon multiple DM-CpG-Loci und wurden als statistisch signifikant erachtet.

Die DM-CpG-Loci von *NR2F1* und *NR2F1-AS1* lagen innerhalb von RDMRs. Der CpG-Locus cg11080549 von *NR2F1-AS1* und die beiden DM-CpG-Loci von *XYLT1* (cg05057352, cg09770579) waren zudem mit einer Enhancer-Region assoziiert (Tabelle 29). cg05057352 (*XYLT1*) lag in einer DHS-Region (DNase I hypersensitive Region) und damit in einer potenziellen Transkriptionsfaktorbindestelle (TFBS). Nach UCSC Genome Bioinformatics

Datenbank und Transkriptionsfaktor-ChIP⁴³-Seq-Track von ENCODE (siehe 2.1.3.2) lag der CpG-Locus cg05057352 von *XYLT1* innerhalb einer Bindestelle des Faktors POLR2A, der größten Untereinheit der Polymerase II (synthetisiert mRNA) (nach NCBI Gene, siehe 2.1.3.2). Der letztgenannte CpG-Locus lag somit innerhalb einer genregulatorischen DNA-Region.

Tabelle 29 DM-CpG-Loci im Herzgewebe von HLHS-Patienten im Vergleich zu fetalem Herzkontrollgewebe von *NR2F1*, *NR2F1-AS1*, *PDXDC1* und *XYLT1*

Gen	CpG ID	Position CHR (GRCh37/hg19)	DHS	DMR	Enhancer-Assoziation	DNA-Methyl.-Diff. HLHS - FetalHerz [$\Delta\beta$ in %]	SD σ_{HLHS}	SD $\sigma_{\text{FetalHerz}}$
<i>NR2F1</i>	cg04192879	chr5:92929473	-	RDMR	-	53	0.08	0.02
<i>NR2F1</i>	cg25449484	chr5:92929372	-	RDMR	-	60	0.09	0.00
<i>NR2F1-AS1</i>	cg11080549	chr5:92909968	-	RDMR	ja	52	0.08	0.00
<i>NR2F1-AS1</i>	cg24579887	chr5:92905693	-	RDMR	-	52	0.06	0.00
<i>XYLT1</i>	cg05057352	chr16:17499310	ja	-	ja	-51	0.06	0.03
<i>XYLT1</i>	cg09770579	chr16:17407430	-	-	ja	-50	0.05	0.03
<i>PDXDC1</i>	cg03245889	chr16:15127173	-	-	-	52	0.02	0.07
<i>PDXDC1</i>	cg06978461	chr16:15127194	-	-	-	60	0.02	0.13

Angaben zu DHS (DNase I hypersensitive Region), DMR (D´differenziell methylierte Region) und Enhancer-Assoziation nach Illumina-Annotation; RDMR: reprogrammierungsspezifische DMR; DNA-Methylierungsdifferenz: AVG-Beta Mittelwerte „HLHS“ (17 Hybridisierungen) minus „FetalHerz“ (fetale Herzgewebeprobe (2 Hybridisierungen)). Positiver Wert = hypermethyliert bei Patienten mit HLHS, negativer Wert = hypomethyliert bei Patienten mit HLHS; SD: Standardabweichung [σ] je für AVG-Beta Mittelwerte von „HLHS“ und „FetalHerz“.

Alle DM-CpG-Loci der vier Gene - *NR2F1*, *NR2F1-AS1*, *PDXDC1* und *XYLT1* – zeigten hohe DNA-Methylierungsdifferenzen von $\Delta\beta > 50\%$ (Tabelle 29 und Abbildung 46). Im wT-Test lag bei den DM-CpG-Loci von *NR2F1* (cg04192879: $p = 1,19 \times 10^{-6}$, cg25449484: $p = 8,39 \times 10^{-15}$), *NR2F1-AS1* (cg11080549: $p = 4,63 \times 10^{-15}$, cg24579887: $p = 2,2 \times 10^{-16}$), und *XYLT1* (cg05057352: $p = 5,4 \times 10^{-4}$, cg09770579: $p = 3,36 \times 10^{-3}$) eine signifikante differenzielle Methylierung zwischen den Gruppen HLHS und fetalem Herzkontrollgewebe vor. Die DM-CpG-Loci von *PDXDC1* zeigte im wT-Test keine signifikanten DNA-Methylierungsdifferenzen mit $p = 0,058$ (cg03245889) und $p = 0,094$ (cg06978461) auf (Grenzwert: $p \leq 0,05$).

⁴³ ChIP-Seq: *Chromatin ImmunoPrecipitation DNA-Sequencing* umfasst eine Methode zur Bestimmung von Protein-DNA-Interaktionen [333]

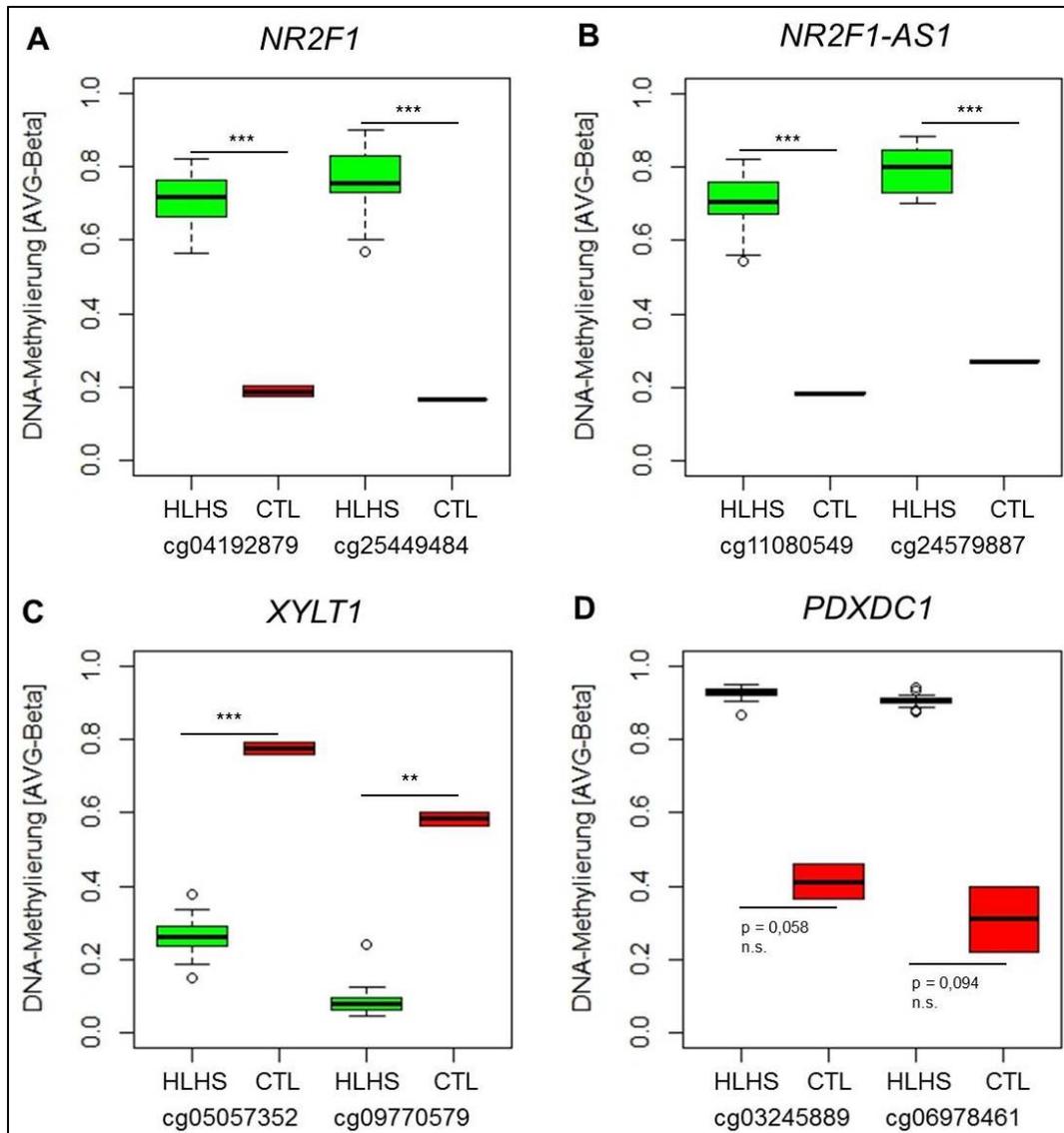


Abbildung 46 Boxplot-Darstellung der DM-CpG-Loci der vier Gene - *NR2F1*, *NR2F1-AS1*, *PDXDC1* und *XYLT1*, welche im Herzgewebe von HLHS-Patienten differenziell methyliert im Vergleich zu fetalem Herzkontrollgewebe (CTL) vorlagen. Die Gruppe „HLHS“ umfasste 17 IAS-Gewebeproben und die Gruppe „CTL“ umfasste 2 fetale Herzgewebeproben. In den Boxplotdarstellungen A – D sind pro Gen und CpG-Locus die DNA-Methylierung (AVG-Beta-Werte) auf der y-Achse und die Probenbezeichnung auf der x-Achse dargestellt. Boxplots der HLHS-Gruppe sind grün und die der CTL-Gruppe rot eingefärbt. Die p-Werte des wT-Tests sind jeweils pro analysiertem CpG nach Sternchencode aufgeführt. Die DM-CpG-Loci von *NR2F1* (A), *NR2F1-AS1* (B) und *XYLT1* (C) zeigten eine signifikante differenzielle Methylierung zwischen den Gruppen „HLHS“ und „CTL“, während die DNA-Methylierungsdifferenzen der CpG-Loci von *PDXDC1* (D) nicht signifikant (n.s.) waren (Grenzwert: $p \leq 0,05$ (*)).

Eine signifikante Anreicherung von RDMR-assoziierten CpG-Loci ($OR = 6,71$, $p = 4,3 \times 10^{-7}$) sowie von Enhancer-assoziierten CpG-Loci ($OR = 5,56$, $p = 7,6 \times 10^{-15}$) konnte innerhalb der 86 „HLHS-spezifischen“ DM-CpG-Loci festgestellt werden. Die 51 assoziierten Gene dieser DM-CpG-Loci zeigten nach GO-Analyse (siehe 2.2.5.3, hier Signifikanzlevel $p \leq 0,05$) eine signifikante **Anreicherung** von Genen im biologischen Prozess „regulation of developmental

process (GO:0050793)“ (11 Gene, $q = 0,0138$), inklusive des Gens *XYLT1*. Darüber hinaus lag eine Anreicherung der molekularen Funktion „protein binding transcription factor activity (GO:0000988)“ (6 Gene, $q = 0,0138$), wozu auch *NR2F1* zählte, vor.

3.2.2.5.3 Verifizierung

Zur Verifizierung der differentiellen DNA-Methylierung im Herzgewebe von HLHS-Patienten verglichen mit fetalem Herzkontrollgewebe wurden jeweils die zwei DM-CpG-Loci der Gene *NR2F1*, *NR2F1-AS1*, *PDXDC1* und *XYLT1* mittels BSPS (Bisulfit-Pyrosequenzierung, siehe 2.2.2.3) analysiert (Primer siehe Abschnitt 8 Tabelle 32). Die Qualitätskontrolle anhand der Konversionskontrollen mithilfe der PyroQ-CpG™ Software (Qiagen) (Details siehe 2.2.2.3 „Qualitätskontrolle“) ergab keine *Peak*-Detektion, sodass von einer effizienten Bisulfit-Konversion ausgegangen werden konnte. Für das HLHS-Kollektiv standen drei Proben (HLHS_13, HLHS_72 und HLHS_130) mit ausreichender DNA-Quantität für die Verifizierung zur Verfügung. Zudem wurden die Proben M_FetalHerz und F_Fetalherz verwendet, sodass die Verifizierung an insgesamt fünf Proben durchgeführt wurde (Rohdaten siehe Abschnitt 8 Abbildung 63 (*NR2F1* cg04192879)).

Zur Analyse der Korrelation der Techniken 450K-Array und BSPS wurden die Methylierungswerte in einem Scatterplot als lineare Regression dargestellt (Abbildung 47). Das Bestimmtheitsmaß R^2 ergab für alle analysierten CpG-Loci Werte nahe 1 (Minimum $R^2 = 0,94$ bei cg05057352), sodass eine hohe Korrelation der gemessenen Methylierungswerte der Techniken 450K-Array und BSPS gegeben war (Details lineare Regressionsanalyse siehe 2.2.4). Die differentielle DNA-Methylierung zwischen den Gruppen HLHS und fetalem Herzkontrollgewebe konnte somit verifiziert werden.

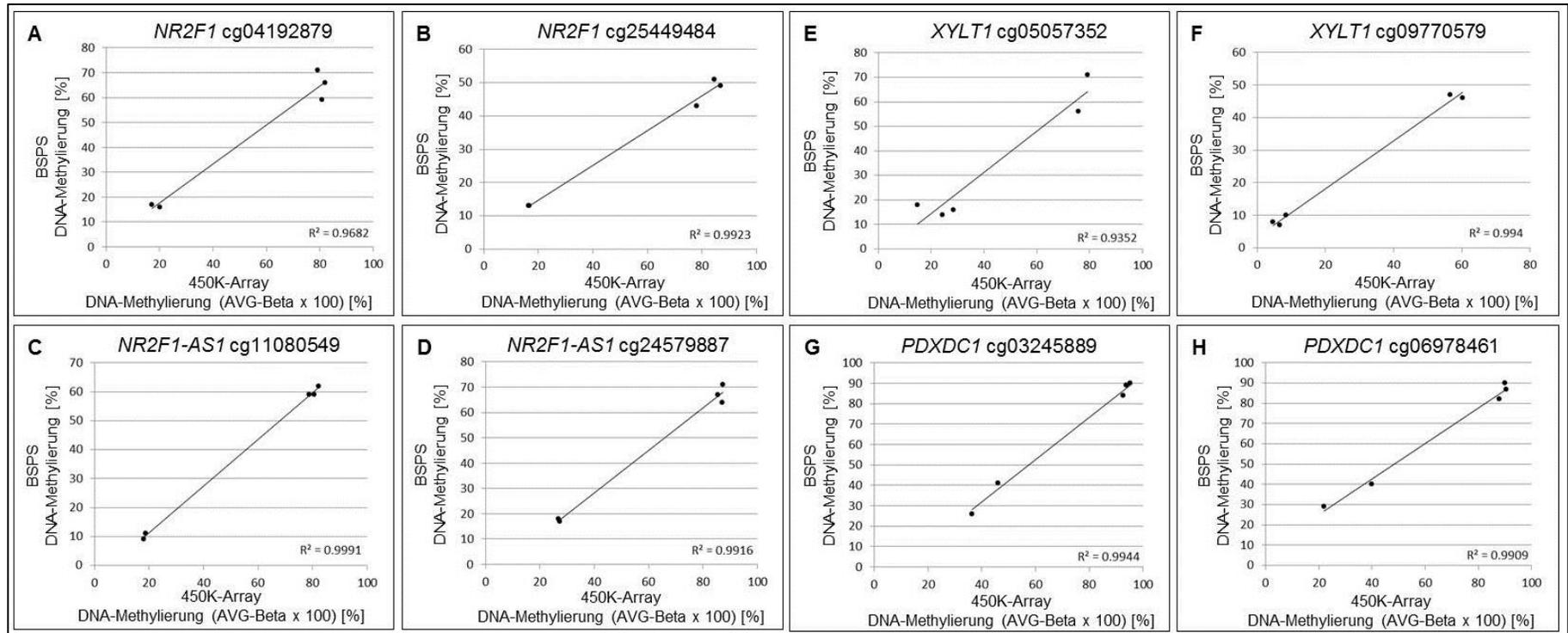


Abbildung 47 Lineare Regression der DM-CpGs der Gene (A) *NR2F1* (A: cg04192879, B: cg25449484), *NR2F1-AS1* (C: cg11080549, D: cg24579887), *XYLT1* (E: cg05057352, F: cg09770579) und *PDXDC1* (G: cg03245889, H: cg06978461) der Korrelationsanalyse (Pearson-Methode) der DNA-Methylierungswerte des 450K-Arrays und der BS-PS. Die aufgeführten CpG-Loci lagen differenziell methyliert zwischen den Gruppen HLHS und fetalem Herzkontrollgewebe in der 450K-Analyse vor. Zur Verifizierung mittels BS-PS wurden drei DNA-Proben des HLHS-Kollektivs (HLHS_13, HLHS_72 und HLHS_130) und die beiden DNA-Proben der fetalen Herzgewebe (M_FetalHerz und F_Fetalherz) analysiert. Methylierungswerte der BS-PS (DNA-Methylierung [%]) auf der y-Achse wurden gegen die 450K-Methylierungswerte (DNA-Methylierung (AVG-Beta x 100) [%]) auf der x-Achse in einem Scatterplot als lineare Regression dargestellt. Das Bestimmtheitsmaß R^2 ergab für cg04192879 $R^2 = 0,97$, für cg25449484 $R^2 = 0,99$, für cg11080549 $R^2 = 0,99$, für cg24579887 $R^2 = 0,99$, für cg05057352 $R^2 = 0,94$, für cg09770579 $R^2 = 0,99$, für cg03245889 $R^2 = 0,99$ und für cg06978461 $R^2 = 0,99$.

Zusammenfassend ergab der Vergleich der DNA-Methylierung von Herzgewebe von HLHS-Patienten und fetalem Herzkontrollgewebe insgesamt 86 „HLHS-spezifische“ DM-CpG-Loci. Hierbei zeigten *NR2F1*, *NR2F1-AS1*, *PDXDC1* und *XYLT1* multiple DM-CpG-Loci, die teilweise mit regulativen DNA-Regionen assoziiert waren. Insgesamt ließen sich hohe Methylierungsdifferenzen von $\Delta\beta > 50\%$ im Herzgewebe von HLHS-Patienten verglichen mit fetalem Herzkontrollgewebe nachweisen.

3.2.3 Identifizierung von 11 differenziell exprimierten Genen bei

HLHS im Vergleich zu Rechtsherzfehlbildungen

Um auf Transkriptomebene Herzgewebeproben von HLHS-Patienten zu analysieren, wurde eine Gesamt-Transkriptom-Sequenzierung (RNA-Sequenzierung (Details siehe 2.2.3.3)) an IAS-Gewebeproben von 25 HLHS-Patienten und 10 Patienten mit Rechtsherzfehlbildungen (Probenzuordnung siehe 2.1.1.1, Tabelle 1 Analyse „R1“ bzw. „R2“) durchgeführt. Hierbei sollte die Fragestellung bearbeitet werden, ob Herzgewebeproben von Patienten mit HLHS eine differenzielle Genexpression im Vergleich zu Herzgewebeproben von Patienten mit einer rechtsseitigen Herzfehlbildung (Rechtsherzfehlbildungen) als Kontrollkollektiv zeigten. Bei der Probenauswahl wurde darauf geachtet, möglichst dieselben IAS-Gewebeproben, die zuvor auch für DNA-Methylierungsanalysen verwendet wurden, für die RNA-Extraktion und nachfolgende RNA-Sequenzierung einzusetzen. Somit lagen zu 23 der 25 analysierten IAS-Gewebeproben von HLHS-Patienten und zu 9/10 IAS-Gewebeproben der Patienten mit Rechtsherzfehlbildungen sowohl DNA-Methylierungsdaten als auch Transkriptom-Daten vor.

Zusätzlich zu den im Vorfeld dieser Arbeit sequenzierten 13 HLHS-Proben und fünf Proben von Patienten mit Rechtsherzfehlbildungen (in Tabelle 1 mit „R1“ gekennzeichnet, prozessiert durch Frau Dr. med. Anne-Karin Kahlert) wurden in der vorliegenden Arbeit für 17 weitere IAS-Gewebeproben von 12 HLHS-Patienten und fünf Patienten mit Rechtsherzfehlbildungen (Tabelle 1 „R2“) die Genexpressionslevel mittels RNA-Sequenzierung bestimmt. Die RNA-Sequenzierung „R1“ hatte eine mittlere *Reads*-Anzahl von 36,8 Mio. *Reads* und „R2“ von 61,4 Mio. *Reads*.

3.2.3.1 Qualitätskontrolle

Zur Qualitätskontrolle wurden die erhaltenen Rohdaten der 17 IAS-Gewebeproben mittels FastQC-Software (Details siehe 2.2.5.4) analysiert. Die Genauigkeit des *Base-Callings* wurde anhand von Phred-Qualitätswerten (Abbildung 48 A) untersucht. Hierbei ergab sich ein Phred-Wert-Median von 36,7 über alle 17 Proben, was einer Wahrscheinlichkeit eines inkorrekten *Base-Callings* von 1 zu 1223 (*Base-Calling*-Genauigkeit > 99,9 %) entsprach. Es konnte somit von einer hohen Qualität des Sequenzierlaufs ausgegangen werden. Des Weiteren wurden Phred-Qualitätswerte pro Sequenz anhand „*per Sequence Quality Score*“-Darstellungen (Details siehe 2.2.5.4) analysiert, um auszuschließen, dass ein größerer Teil der Sequenzen geringere Phred-Qualitätswerte aufwies. Hierbei ergab sich ein mittlerer Phred-Wert pro Read von > 37, was einer Fehlerrate deutlich kleiner dem Grenzwert von 0,2 % (Phred = 27) entsprach (Abbildung 48 B). Zudem war ein einzelner *Peak* der mittleren Sequenz-Qualität – und nicht weitere *Peaks* – erkennbar. Somit konnte ausgeschlossen werden, dass es einen Anteil von Sequenzen mit geringerer Sequenz-Qualität gab. Auch die visuelle Analyse der Graphiken zum GC-Gehalt pro Sequenz (*per Sequence GC Content*) ergab eine gute Übereinstimmung mit der modellierten Normalverteilung bei allen 17 IAS-Gewebeproben (Beispiel HLHS_128 in Abbildung 48 C). Der Anteil abweichender *Reads* von der theoretischen Verteilungskurve war kleiner 15 % („*pass filter*“ *Threshold* der FastQC-Software), was für eine gute randomisierte *cDNA-Library* sprach. In einer normalen randomisierten *cDNA-Library* würde eine annähernde Normalverteilung des GC-Gehalts mit einem zentral gelegenen *Peak* beim durchschnittlichen GC-Gehalt des zugrundeliegenden Transkriptoms (berechnet durch die FastQC-Software) erwartet werden [334]. Die Kurvenform des GC-Gehalts der beispielhaft dargestellten Probe HLHS_128 (Abbildung 48 C) stimmte mit der theoretischen Verteilungskurve überein. Kontaminationen durch Adapter-Sequenzen (*Reads* enthalten partielle oder komplette Adaptersequenzen) wären in diesen Graphiken zum GC-Gehalt pro Sequenz erkennbar, indem ein bis mehrere extreme (schmale hohe) *Peaks* auftreten würden und die Kurve nicht der theoretischen Verteilungskurve entspräche. Solche Kontaminationen lagen demnach in dem hier analysierten Datensatz nicht vor.

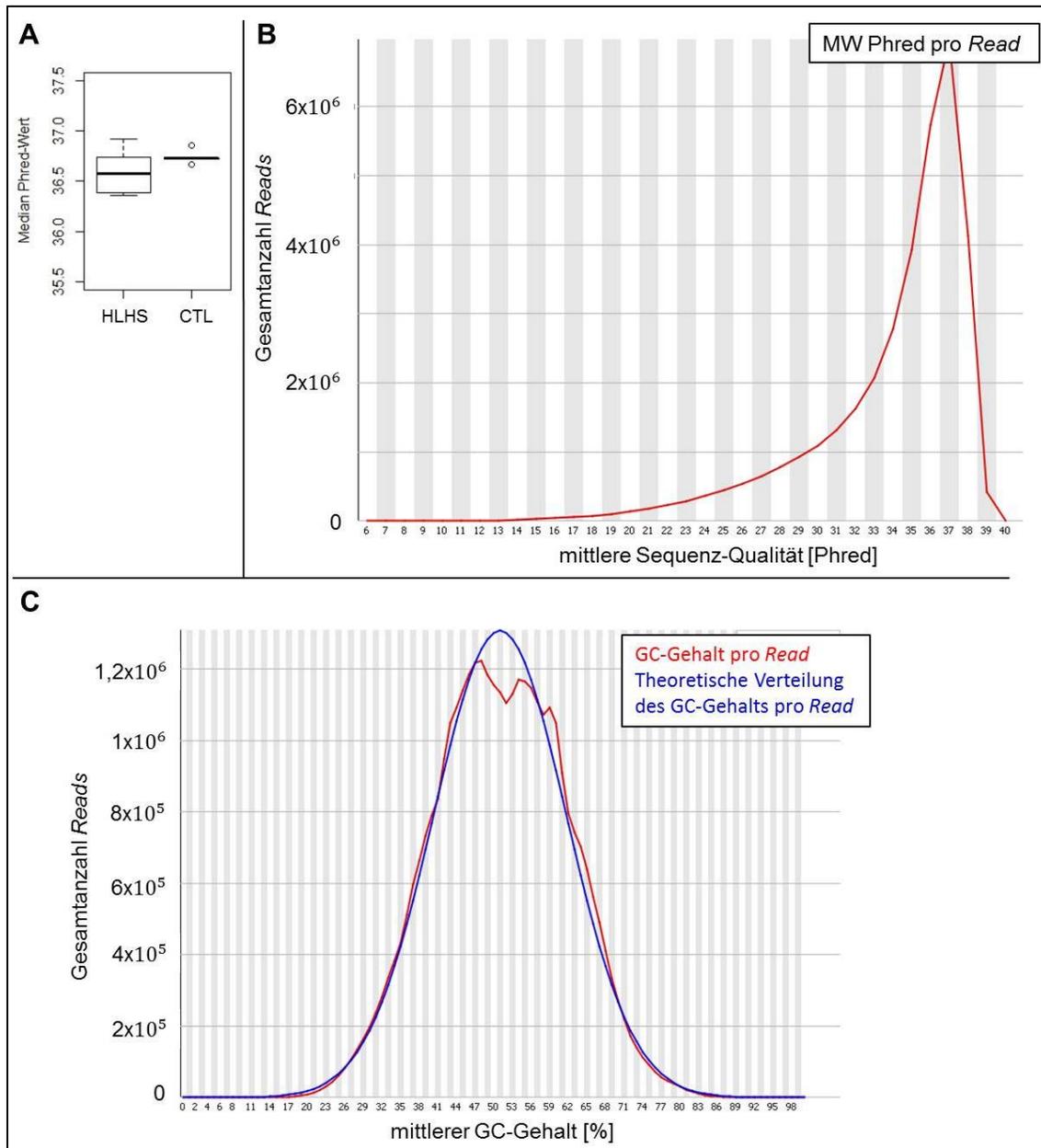


Abbildung 48 Qualitätskontrolle von Rohdaten der RNA-Sequenzierung von 17 IAS-Gewebeproben (12 HLHS-Proben und fünf Proben von Patienten mit Rechtsherzfehlbildungen) mittels FastQC-Software. A: Boxplot Phred-Wert-Mediane von HLHS- und Rechtsherzfehlbildung-Proben („CTL“) mit einem Phred-Wert-Median von 36,7 über alle 17 Proben. Dies entspricht einer *Base-Calling*-Genauigkeit > 99,9 %. B: „per Sequence Quality Score“-Darstellung beispielhaft anhand HLHS_128. x-Achse: mittlere Sequenz-Qualität [Phred], y-Achse: Gesamtanzahl aller Reads. Der mittlere Phred-Wert pro Read ist bei der Probe HLHS_128 Phred > 37. C: „per Sequence GC Content“-Darstellung: GC-Gehalt über jeden Sequenz-Read (rote Kurve) im Vergleich zur modellierten Normalverteilung der GC-Werte (blaue Kurve) beispielhaft anhand HLHS_128. x-Achse: mittlerer GC-Gehalt [%], y-Achse: Gesamtanzahl aller Reads.

3.2.3.2 Datenanalyse

Die bioinformatische Auswertung, in der eine differenzielle **Expressionsanalyse** mittels EdgeR-Software [254] durchgeführt wurde, erfolgte durch PhD Gabriel Musso (Brigham and Women's

Hospital, Harvard Medical School, Boston). Die Analyse der beiden Sequenzierläufe (Tabelle 1, „R1“ und „R2“), mit insgesamt 26 IAS-Gewebeproben von HLHS-Patienten und 10 IAS-Gewebeproben von Patienten mit Rechtsherzfehlbildungen, ergab im EdgeR *Fisher's Exact Test* ($p \leq 0,05$) bei HLHS-Patienten verglichen mit Patienten mit Rechtsherzfehlbildungen eine signifikante Hochregulation der Expression von neun Genen: *ACE2*, *HOGA1*, *MLLT11*, *CASQ1*, *CSPG4*, *SERPINE2*, *XYLT1*, *ATP10A* und *PDGFRA*. Zudem konnten zwei signifikant herunterregulierte Gene – *SOSTDC1* und *C1orf115* – im IAS-Gewebe von HLHS-Patienten im Vergleich zu Patienten mit Rechtsherzfehlbildungen identifiziert werden. Das Gennetzwerk (GeneMANIA [263], siehe 2.1.3.2) dieser vorwiegend Zellzyklus-assoziierten differenziell exprimierten Gene ist in Abbildung 49 dargestellt.

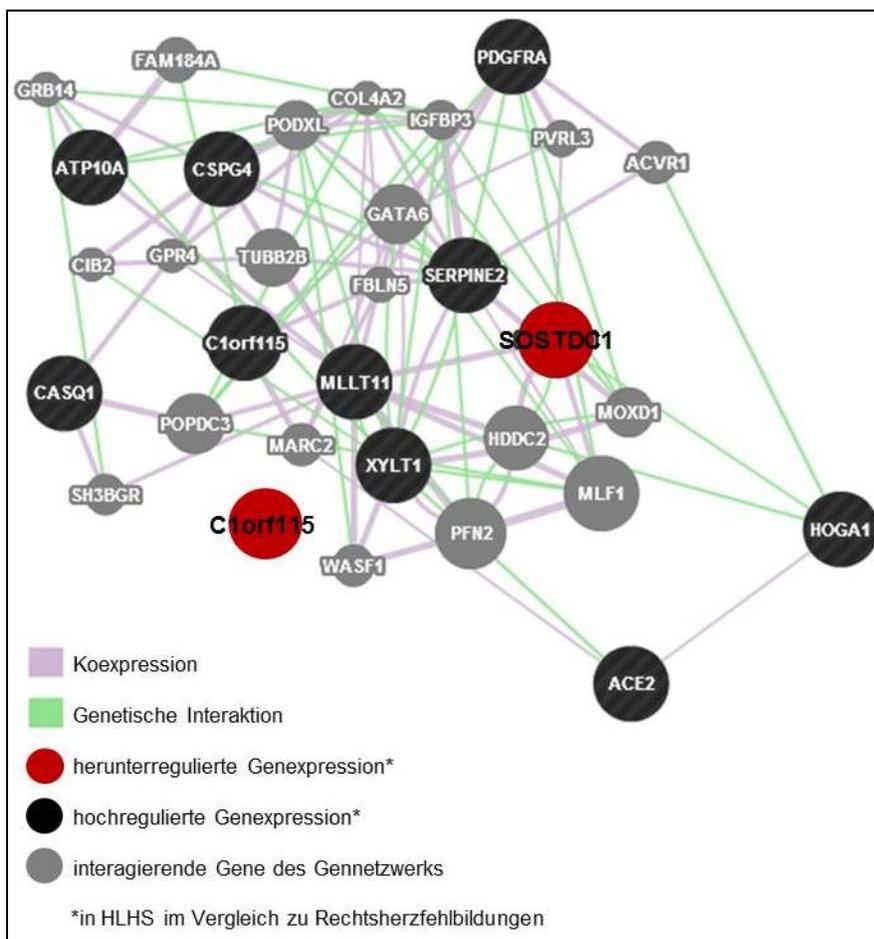


Abbildung 49 Gennetzwerk nach GeneMANIA der differenziell exprimierten Gene im IAS-Gewebe von 26 HLHS-Patienten und 10 Patienten mit Rechtsherzfehlbildungen. Im EdgeR *Fisher's Exact Test* lag eine signifikante Hochregulation der Genexpression im IAS-Gewebe von HLHS-Patienten im Vergleich zu Patienten mit Rechtsherzfehlbildungen von folgenden Genen vor (schwarze Füllung): *ACE2*, *HOGA1*, *MLLT11*, *CASQ1*, *CSPG4*, *SERPINE2*, *XYLT1*, *ATP10A* und *PDGFRA*. Die Gene *SOSTDC1* und *C1orf115* waren signifikant herunterreguliert im IAS-Gewebe von HLHS-Patienten im Vergleich zu Patienten mit Rechtsherzfehlbildungen. Interagierende Gene des Gennetzwerks sind durch eine graue Füllung dargestellt. Die Koexpression der Gene ist durch violette Linien und die genetische Interaktion durch grüne Linien gekennzeichnet.

3.2.3.3 Verifizierung

Das Gen *PDGFRA* zeigte nach MGI-Datenbank bei *PDGFRA*-Knockout-Mäusen einen kardiovaskulären Phänotyp sowie Auswirkungen auf die Embryogenese und wurde daher näher analysiert. Mittels quantitativer Real-Time PCR (qPCR; Details siehe 2.2.3.5) erfolgte die

Verifizierung mit drei HLHS-Proben und drei Proben von Patienten mit Rechtsherzfehlbildungen (hier TA-Proben). Lediglich bei diesen Proben war restliche RNA zur Generierung von cDNA für die qPCR-Analyse vorhanden. Letztere Analyse ergab jedoch keine Bestätigung der Hochregulation von *PDGFRA* bei HLHS im Vergleich zu Rechtsherzfehlbildungen (Abbildung 50). Innerhalb letztgenannter Proben lag eine starke Heterogenität der *Ratio*-Werte der mRNA-Expression vor (Abbildung 50 B und C). Der *Ratio*-Wert der HLHS-Proben lag mit einem Median von 4,06 höher als der der TA-Proben mit einem Median von 3,7 (Abbildung 50 C). Dies entsprach tendenziell der nach EdgeR-Analyse ermittelten Hochregulation der *PDGFRA*-Expression bei HLHS, jedoch wies letztere Gruppe eine hohe Streuung auf (Abbildung 50 B).

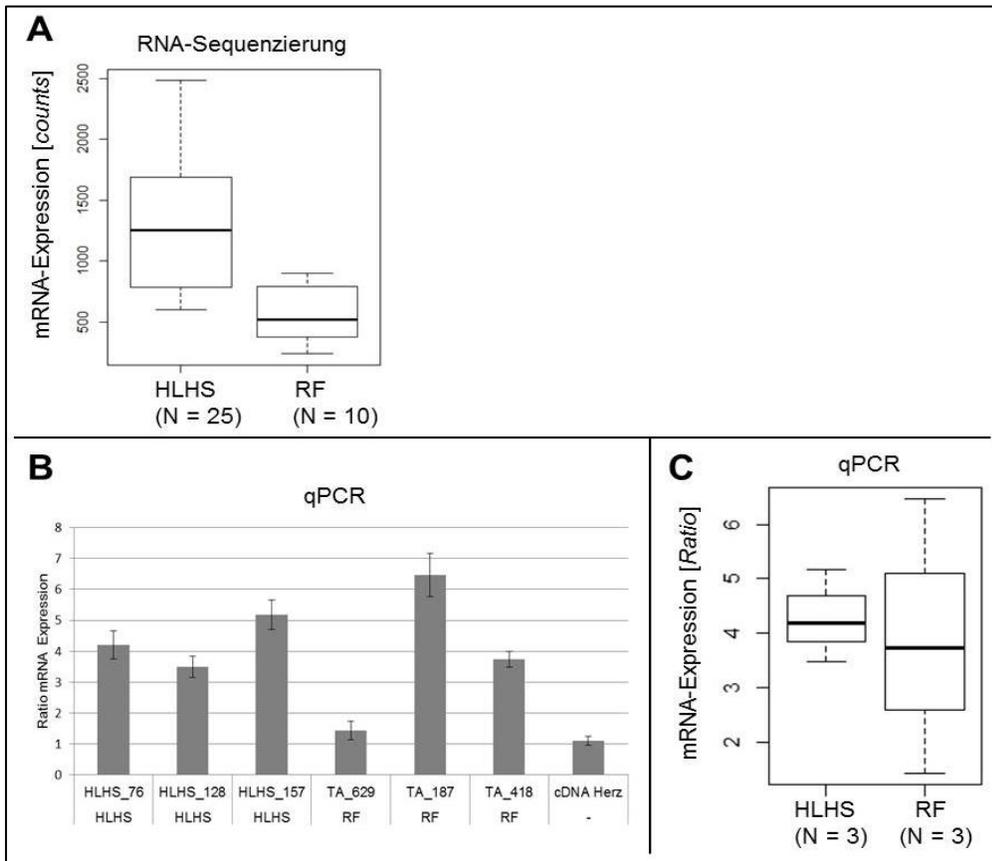


Abbildung 50 Ergebnis der qPCR-Analyse von Proben von HLHS-Patienten und Patienten mit Rechtsherzfehlbildungen zur Verifizierung von Genexpressionsdaten zu *PDGFRA* aus der RNA-Sequenzierungsanalyse. A: Boxplot-Darstellung der mRNA-Expression in *counts* (Anzahl gemappter *Reads*) der mittels RNA-Sequenzierung analysierten Proben von HLHS-Patienten und Patienten mit Rechtsherzfehlbildungen (RF). Eine hochregulierte Genexpression von *PDGFRA* bei HLHS im Vergleich zu Rechtsherzfehlbildungen ist im Boxplot zu erkennen. B: Balkendiagramm der *Ratio*-Mittelwerte der mRNA-Expression (mit Standardabweichung) aus der qPCR-Analyse von drei unabhängigen technischen cDNA-Replikaten, hergestellt aus RNA von HLHS- und RF-Proben. x-Achse: Probenbezeichnung der cDNA-Proben, y-Achse: *Ratio*-Werte der mRNA-Expression. Herz-cDNA (Clontech Laboratories Inc. (Mountain View, USA)) wurde als Positivkontrolle verwendet. Es lag eine hohe Intragruppen-Heterogenität vor. C: Boxplot-Darstellung der *Ratio*-Mittelwerte der mittels qPCR analysierten HLHS- und Rechtsherzfehlbildungen-Proben. Anhand der qPCR-Analyse ist kein signifikanter Unterschied der Genexpression zwischen den Gruppen HLHS und Rechtsherzfehlbildungen festzustellen.

Aufgrund des limitierten RNA-Materials konnten keine weiteren qPCR-Analysen durchgeführt werden, um die Heterogenität der Expressionswerte genauer zu analysieren. Es stand von keinen weiteren Proben der RNA-Sequenzierungsanalyse restliche RNA zur Verfügung, sodass mögliche Degradationseffekte der RNA (Lagerung teilweise > 1 Jahr) und dadurch potenziell hervorgerufene Varianzen der *Ratio*-Werte der mittels qPCR analysierten Proben nicht nähergehend untersucht werden konnten. Da es sich jedoch um schwer zugängliches primäres Herzgewebe handelt, wurden die durch PhD Gabriel Musso (Brigham and Women's Hospital, Harvard Medical School, Boston) zur Verfügung gestellten Daten zur differenziellen Expression bei HLHS im Vergleich zu Rechtsherzfehlbildungen im Folgenden für integrative Analysen weiter verwendet. Darüber hinaus fanden die Ergebnisse des *SNP-Callings* aus der RNA-Sequenzierung im Abgleich mit WES-Daten Anwendung.

3.3 Integrative Analysen von Genom, Epigenom und Transkriptom

Um die molekularen Mechanismen in der Pathogenese des HLHS integrativ zu untersuchen, wurden die Ergebnisse der genetischen Analysen auf Keimbahn- (siehe 3.1) und somatischer Ebene (3.2), der DNA-Methylierungsanalysen (3.2.2) und der Transkriptom-Analysen (3.2.3) gemeinsam analysiert. Hierbei wurden die Datensätze hinsichtlich überlappender Gene mit Mutationen, differenzieller DNA-Methylierung und/oder differenzieller Expression untersucht. Mögliche Überlappungen betroffener *Gen-Pathways* auf Genom-, Epigenom- und Transkriptomebene sollten so identifiziert werden.

3.3.1 Identifizierung von 20 differenziell methylierten Genen mit zusätzlichen funktionellen Genvarianten bei Patienten mit CHDs

In der folgenden integrativen Analyse wurden die in dieser Arbeit bei CHD-Patienten identifizierten differenziell methylierten Gene (siehe 3.2.2) mit detektierten funktionellen Genvarianten verglichen. Hierbei wurden sowohl die in Abschnitt 3.1 auf **Keimbahnebene**, als

auch die in Abschnitt 3.2.1 auf **somatischer Ebene** (im Herzgewebe) detektierten Genvarianten abgeglichen.

Bei dem Gen *NR2F2* (NM_021005.3) konnte sowohl eine differenzielle Methylierung multipler CpG-Loci in der Herz-Mapping-Analyse (siehe 3.2.2.2) als auch eine kausale *de novo* Splice-Donor-Mutation [303] in der **Keimbahn** eines HLHS-Patienten nachgewiesen werden (siehe 3.1.2.1). In verschiedenen Herzregionen von CHD-Patienten zeigte das Gen *NR2F2* fünf DM-CpG-Loci (siehe 3.2.2.2 Tabelle 28) deren AVG-Beta-Mittelwerte verschiedener Herzgewebetypen in Abbildung 51 dargestellt ist. Die RDMR-assozierten CpG-Loci zeigten hierbei vorwiegend eine Hypomethylierung in ventrikulärem, atrialem und fetalem Herzgewebe im Vergleich zu einer tendenziellen Hypermethylierung von Herz-Klappengewebe und Gefäß-Strukturen. Die RDMRs lagen vorwiegend in Assoziation zu CpG-Insel-Shores vor, d.h. 5'/3'-angrenzende Bereiche der CpG-Insel [335].

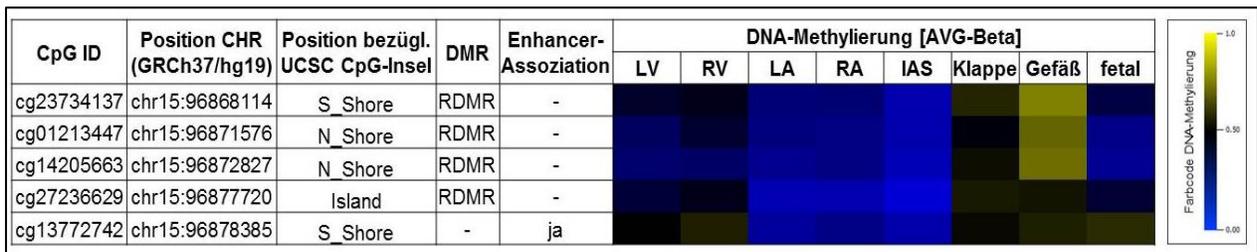


Abbildung 51 DM-CpG-Loci von *NR2F2* der Herz-Mapping-Analyse (siehe 3.2.2.2). Die DNA-Methylierung der CpG-Loci von *NR2F2* ist als AVG-Beta-Mittelwert anhand definierter Herzgewebetypen (LV, RV, ...), beschrieben in Abschnitt 3.2.2.2, dargestellt. Der Farbcode der DNA-Methylierung ist rechts dargestellt. Zu jedem CpG-Locus (CpG ID) sind die chromosomale Position (Position CHR), die Position im Genom bezüglich bekannter UCSC CpG-Inseln, das Vorliegen einer Überlappung zu einer DMR sowie eine Enhancer-Assoziation nach Illumina-Annotation aufgeführt. Island: CpG-Insel; N_Shore: North_Shore (5' angrenzender Bereich zur CpG-Insel); S_Shore: South Shore (3' angrenzender Bereich zur CpG-Insel); fetal: fetales Herzkontrollgewebe.

Auf **somatischer Ebene** erfolgte in dieser integrativen Analyse ein Abgleich differenziell methylierter Gene aus T-Test-Analysen des Abschnitts 3.2.2 mit Genen, in denen mittels WES-Analyse von IAS-Gewebe von HLHS-Patienten funktionelle Genvarianten identifiziert wurden (siehe 3.2.1.1). In 16 Genen mit funktionellen Genvarianten im Herzgewebe von Patienten mit HLHS konnte zusätzlich eine differenzielle Methylierung in Herzmuskelgewebe verglichen mit Gefäßstrukturen in der Muskel-Gefäß-Analyse (siehe 3.2.2.4) nachgewiesen werden. Hierbei ließen sich > 50 %ige Methylierungsdifferenzen feststellen, welche in Abbildung 64 (siehe Abschnitt 8) aufgeführt sind.

Hierbei ist das Gen *DOCK1* (NM_001380.3) hervorzuheben, welches zusätzlich zu einer Hypermethylierung eines CpG-Locus in „Muskel“ im Vergleich zu „Gefäß“ eine *compound*

heterozygote Mutation beim Patienten HLHS_132 zeigte (Abbildung 52). Die Mutation lag sowohl im initial mittels WES-Technik analysierten Herzgewebe von HLHS_132 als auch in der Keimbahn vor. Es handelte sich um zwei bisher unbekannte Missense-Mutationen, die beim Indexpatienten *compound* heterozygot auftraten und jeweils von einem Elternteil vererbt wurden: die maternal vererbte Mutation c.1927G>A (p.Asp643Asn) und die paternal vererbte Mutation c.5038A>G (p.Arg1680Gly). Diese Mutationen werden nach Polyphen-2 als *probably* bzw. *possibly damaging* eingestuft und konnten mittels Sanger-Sequenzierung in der Familie 6 nachgewiesen werden (Abbildung 52 A).

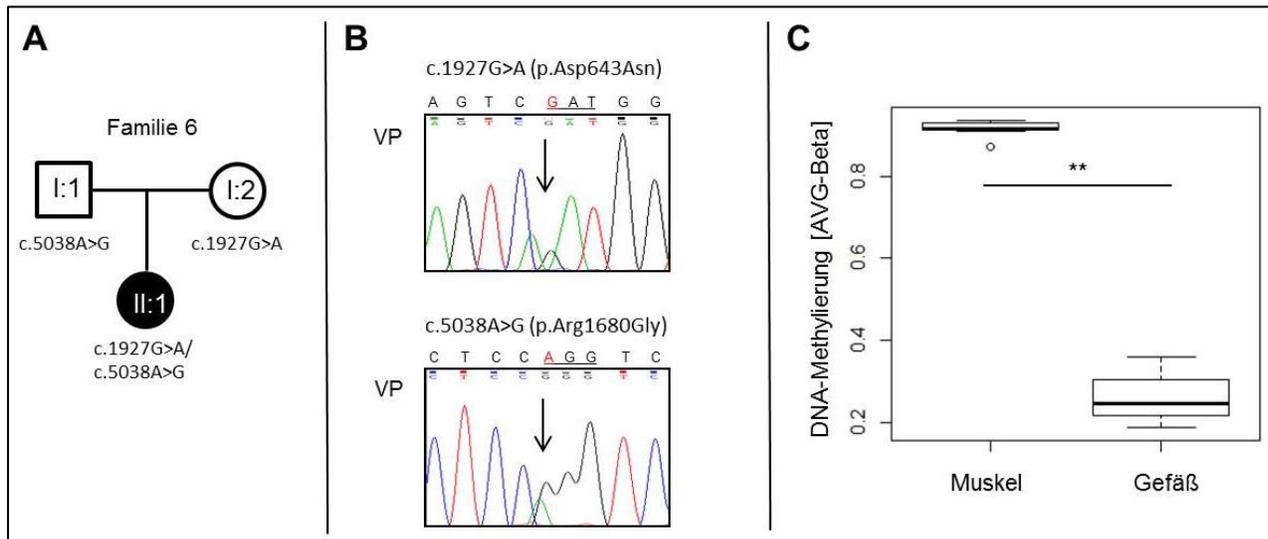


Abbildung 52 *Compound* heterozygote Mutationen in *DOCK1* bei einem HLHS-Patient und differenzielle DNA-Methylierung von *DOCK1* in „Muskel“ und „Gefäß“ A: Stammbaum mit Genotyp der Mutationen in *DOCK1* der Familie 6. Die phänotypisch gesunden Eltern sind in weißer Füllung mit I:1 Vater (heterozygot c.5038A>G) und I:2 Mutter (c.1927G>A) dargestellt. Beim Indexpatient HLHS_132 (II:1, schwarze Füllung) konnte eine *compound* heterozygote Mutation (c.1927G>A/c.5038A>G) nachgewiesen werden. B: Elektropherogramme der mittels Sanger-Sequenzierung verifizierten Missense-Mutationen in *DOCK1* (c.1927G>A (p.Asp643Asn), c.5038A>G (p.Arg1680Gly)) bei HLHS_132. Dargestellt ist jeweils die Sequenz des Vorwärts-Primers (VP). Rote Buchstaben stellen die mutierte Base dar, unterstrichene Buchstaben die mutierte Kodonsequenz der Aminosäuren. C: Boxplot-Darstellung der differenziellen DNA-Methylierung in AVG-Beta-Werten der Gewebegruppen „Muskel“ (LA, RA, IAS, LV, RV, TV, MV) vs. „Gefäß“ (PA, Aorta, AoV). Es lag im wT-Test eine signifikante ($p = 0,00514$) Hypermethylierung in „Muskel“ im Vergleich zu „Gefäß“ vor.

Die Gene *CTNBN1*, *STEAP2* und *PDXDC1* (Abbildung 53) wiesen im IAS-Gewebe von Patienten mit HLHS seltene bzw. bisher unbekannte als *probably damaging* eingestufte Genvarianten auf (WES-Analyse, siehe 3.2.1.1). Zusätzlich zeigten diese Gene eine signifikant differenzielle Methylierung in demselben Gewebe dieser Patienten verglichen mit fetalem Herzkontrollgewebe mit 50%igen Methylierungsdifferenzen (HLHS-Fetal-Analyse, siehe 3.2.2.5).

Exom						Epigenom		
Fall	Gen	Mutation Position CHR (GRCh37/hg19)	Typ	PPH	ExAC AF	CpG ID Position CHR (GRCh37/hg19)	DNA-Methylierung [AVG-Beta]	
							HLHS	Fetal
HLHS_21/22	<i>CTNNB1</i>	c.1694G>A chr3:41277225	Miss.	D	8.99E-05	cg09678212 chr3:41240163		
HLHS_15	<i>PDXDC1</i>	c.35C>T chr16:15091623	Miss.	D	6.27E-05	cg03245889 chr16:15127172		
						cg06978461 chr16:15127193		
HLHS_12	<i>STEAP2</i>	c.276delT chr7:89854672	FS	-	-	cg02206373 chr7:89841767		

Abbildung 53 Mutationen und DM-CpG-Loci zu drei überlappenden Genen aus der WES-Analyse von IAS-Gewebe von neun HLHS-Patienten (siehe 3.2.1.1) und der DNA-Methylierungsanalyse „HLHS-Fetal“ (siehe 3.2.2.5). Zu jedem HLHS-Fall (Spalte ganz links) sind Genname, Genotyp der Mutation und chromosomale Position (Mutation Position CHR), Mutationstyp (Typ), prädiktive Mutationskonsequenz nach Polyphen-2 (PPH), Allelfrequenz (AF) nach ExAC (Version 0.3, Feb. 2015) und überlappende DM-CpG-Loci (CpG ID) mit chromosomaler Position des CpG-Locus dargestellt. Die DNA-Methylierung (AVG-Beta-Mittelwerte) der Gruppen „HLHS“ (IAS-Gewebe) und „Fetal“ (fetales Herzgewebe) ist rechts aufgeführt. Der Farbcode entspricht der Legende rechts der Abbildung. Dargestellte AVG-Beta-Mittelwerte stimmen mit Einzelwerten des jeweiligen HLHS-Falls überein. Miss.: Missense-Mutation, FS: Frameshift-Mutation. B = *benign*, P = *possibly damaging*, D = *probably damaging*. E = Zehnerpotenz.

Das Gen *CTNNB1* (NM_001904.3) zeigte in den Patienten HLHS_21 und HLHS_22 (monozygote Zwillinge) eine seltene Missense-Mutation (c.1694G>A (p.R565H) ExAC-AF = 8.992×10^{-5}), nach Polyphen-2 als *probably damaging* eingestuft, die durch Sanger-Sequenzierung verifiziert werden konnte. Zusätzlich konnte in diesem Gen eine Hypermethylierung in demselben analysierten IAS-Gewebe dieser HLHS-Patienten verglichen mit fetalem Herzkontrollgewebe detektiert werden (Abbildung 53). *CTNNB1* stellt zudem ein Netzwerk-Interaktionspartner eines Interaktionsnetzwerks aus Entwicklungsprozess-assoziierten Proteinen dar, das in Notch-Signalwegen sowie in der transkriptionellen Regulation involviert ist (siehe STRING-Analyse in Abschnitt 3.2.1.1). Acht Gene dieses Netzwerks wiesen funktionelle Mutationen im IAS-Gewebe von Patienten mit HLHS auf (3.2.1.1, Abbildung 23).

In dem Gen *PDXDC1* (NM_015027) konnte eine seltene Missense-Mutation (c.35C>T (p.T12M), ExAC-AF = 6.266×10^{-5}) bei HLHS_15, die nach Polyphen-2 als *probably damaging* eingestuft wird, nachgewiesen werden. Darüber hinaus zeigte dieses Gen zwei DM-CpG-Loci mit einer Hypermethylierung im IAS-Gewebe dieses Patienten im Vergleich zu fetalem Herzkontrollgewebe.

In *STEAP2* (NM_001040666.1) konnte eine bisher unbekannte Frameshift-Mutation mit vorzeitigem Stopkodon (c.276delT (p.I93Yfs*12)) bei HLHS_12 detektiert und durch Sanger-Sequenzierung verifiziert werden (siehe 3.2.1.1.2 Tabelle 23). Zusätzlich konnte in diesem Gen eine Hypomethylierung im IAS-Gewebe dieses HLHS-Patienten verglichen mit fetalem

Herzkontrollgewebe detektiert werden (Abbildung 53). Die differenzielle DNA-Methylierung der DM-CpG-Loci der Gene *CTNNB1*, *PDXDC1* und *STEAP2* konnte mittels BSPS mit einer hohen Korrelation der Daten ($R^2 > 0,99$) verifiziert werden (*PDXDC1* siehe 3.2.2.5 Abbildung 47 (lineare Regression); *CTNNB1* und *STEAP2* siehe Anhang (Abschnitt 8 Abbildung 65 (lineare Regression) und Abbildung 66 (Rohdaten)).

Zusammenfassend konnten in 20 Genen seltene bzw. bisher unbekannte funktionelle Genvarianten, teilweise als *probably damaging* eingestuft, bei Patienten mit HLHS nachgewiesen werden. Diese Gene zeigten im Herzgewebe von Patienten mit unterschiedlichen CHDs, inklusive dem HLHS, zusätzlich eine differenzielle Methylierung mit $\Delta\beta \geq 50\%$ bei CpG-Loci von teils regulativen DNA-Regionen.

3.3.2 Identifizierung von *ACE2* und *SOSTDC1* mit einer differenziellen Genexpression und funktionellen Genvarianten bei Patienten mit HLHS

In dieser integrativen Analyse sollte untersucht werden, ob Gene mit detektierten funktionellen Genvarianten bei Patienten mit HLHS (HLHS-WES-Datensatz der WES-Analyse, siehe 3.2.1.1) zusätzlich eine aberrante Genexpression im analysierten IAS-Gewebe dieser Patienten aufwiesen. Dazu wurde der HLHS-WES-Datensatz (hier $AF \leq 5\%$ als Filterkriterium) mit den, in demselben Gewebe identifizierten, differenziell exprimierten Genen (RNA-Sequenzierung, siehe 3.2.3) abgeglichen. Die Gene *ACE2* (NM_021804.2) und *SOSTDC1* (NM_015464.2) zeigten hierbei sowohl auf Transkriptomebene eine differenzielle Expression als auch auf Genomebene proteinverändernde Mutationen.

Das Gen *ACE2* zeigte bei HLHS_4 eine bisher unbekannte Missense-Mutation (c.2188A>C (p.T730P)), nach Polyphen-2 als *probably damaging* eingestuft, und lag im Herzgewebe von HLHS-Patienten im Vergleich zu Patienten mit Rechtsherzfehlbildungen signifikant hochreguliert ($p = 0,003857$) vor (Abbildung 54). Die Missense-Mutation konnte mittels Sanger-Sequenzierung in hemizygoter Ausprägung beim Indexpatienten HLHS_4 und in heterozygoter Ausprägung bei der Mutter nachgewiesen werden (Abbildung 54 B). Leider konnte von dem gesunden

Geschwisterkind (Abbildung 54 C, II:2) keine pB-Probe erhalten werden, um bei ihm ein Auftreten der Mutation auszuschließen.

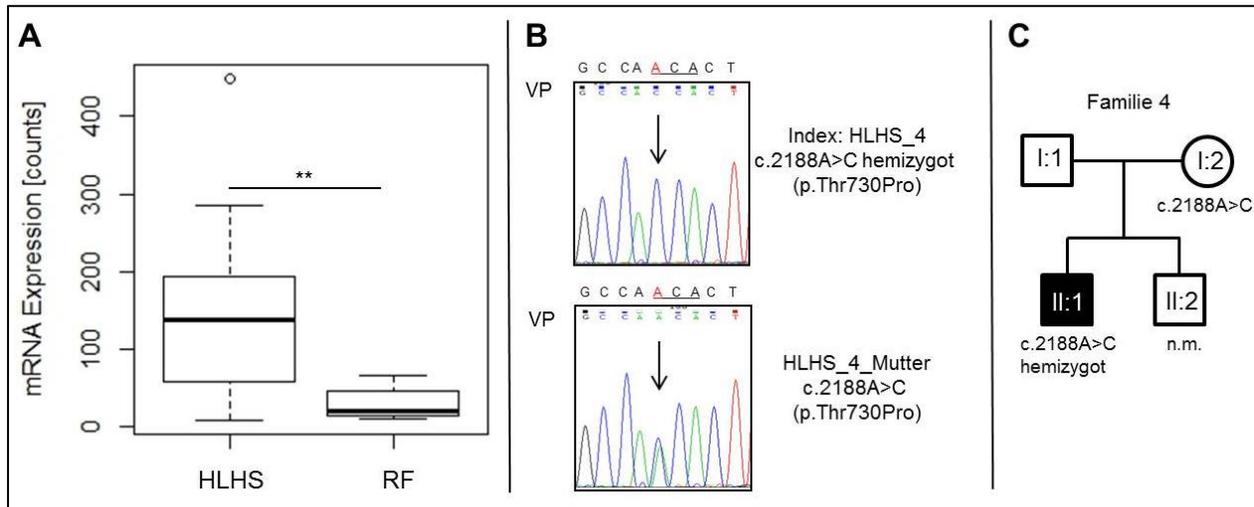


Abbildung 54 Differenzielle Expression und funktionelle Genvariante in *ACE2* (NM 021804.2). A: Boxplot-Darstellung der differentiellen Genexpression in *ACE2* im IAS-Gewebe von HLHS (N = 26) im Vergleich zu Rechtsherzfehlbildungen (RF) (N = 10). Die mRNA-Expression ist in *counts* (Anzahl gemappter Reads) auf der y-Achse angegeben. Nlm *Fisher's Exact* Test lag eine signifikant hochregulierte Expression von *ACE2* in „HLHS“ im Vergleich zu „RF“ mit $p = 0,003857$ vor (Grenzwert: $p \leq 0,05$). B: Elektropherogramm der mittels Sanger-Sequenzierung verifizierten Missense-Mutation c.2188A>C (p.T730P) in *ACE2*: hemizygot bei HLHS_4 (im IAS-Gewebe und in pB detektiert) und heterozygot bei HLHS_4_Mutter. Dargestellt ist jeweils die Sequenz des Vorwärts-Primers (VP). Rote Buchstaben stellen die mutierte Base dar, unterstrichene Buchstaben die mutierte Kodonsequenz der Aminosäuren. C: Stammbaum der Familie 4 (Tabelle 2) mit Genotyp der Mutation in *ACE2*. Eine schwarze Füllung symbolisiert den Indexpatienten mit HLHS, während eine weiße Füllung für nicht betroffene Familienmitglieder steht. Der Vater (I:1) wies keine Mutation in *ACE2* auf, die Mutter (I:2) zeigte eine heterozygote Ausprägung der Mutation (c.2188A>C) und der Indexpatient mit HLHS (II:1) zeigte eine hemizygoten Ausprägung der Mutation (c.2188A>C hemizygot). Vom gesunden Bruder (II:2) konnte keine pB-Probe erhalten werden (n.m. = nicht möglich).

Das Gen *SOSTDC1* zeigte bei HLHS_132 eine Frameshift-Mutation c.592delA (p.S198Afs*8 (ExAC-AF = 0,34 %)), die nach Polyphen-2 als *probably damaging* eingestuft wird, und lag im Herzgewebe von HLHS-Patienten im Vergleich zu Patienten mit Rechtsherzfehlbildungen herunterreguliert vor. Zur Untersuchung der herunterregulierten Expression dieses Gens wurde ein Knockdown im Zebrafisch-Modell mittels Morpholino-Injektion (Sequenz-spezifische Translations-Inhibitoren) durch die Kollegin Frau Dr. med. Anne-Karin Kahlert durchgeführt (Experimente erfolgten am Brigham and Women's Hospital (Harvard Medical School, Boston)). Die Analysen im Zebrafisch-Modell ergaben eine kardiale Dysfunktion (kardiales Ödem) und muskuläre Defekte am Rückgrat (Daten wurden noch nicht publiziert). In Abbildung 55 ist das Elektropherogramm der identifizierten Frameshift-Mutation und die Boxplot-Darstellung der differentiellen Genexpression von *SOSTDC1* dargestellt. Im *Fisher's Exact* Test war diese nicht signifikant, tendenziell war jedoch das *SOSTDC1*-Gen im Herzgewebe von HLHS-Patienten verglichen mit Patienten mit Rechtsherzfehlbildungen geringer exprimiert (Abbildung 55 A). Die

Frameshift-Mutation c.592delA (p.S198Afs*8) lag sowohl beim Indexpatienten HLHS_132 als auch beim phänotypisch gesunden Vater heterozygot vor (Abbildung 55 C).

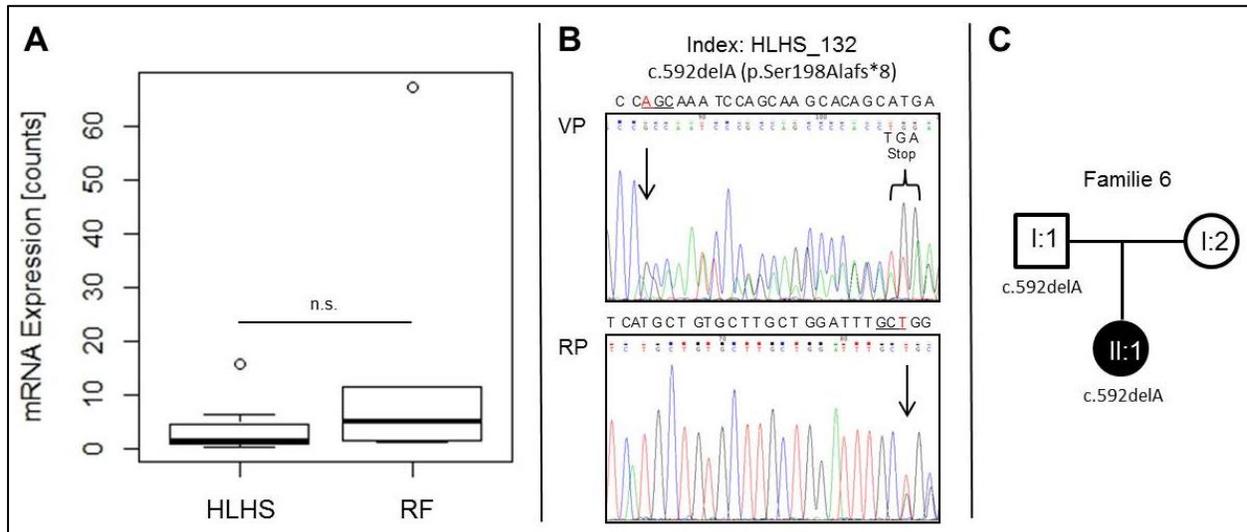


Abbildung 55 Differenzielle Expression und funktionelle Genvariante in *SOSTDC1* (NM 015464.2). A: Boxplot-Darstellung der differentiellen Genexpression in *SOSTDC1* im IAS-Gewebe von HLHS (N = 26) im Vergleich zu Patienten mit Rechtsherzfehlbildungen (RF) (N = 10). Die mRNA Expression ist in counts (Anzahl gemappter Reads) auf der y-Achse angegeben. Im Fisher's Exact Test lag keine signifikante differenzielle Expression zwischen den Gruppen „HLHS“ und „RF“ vor (Grenzwert: $p \leq 0,05$), tendenziell war jedoch *SOSTDC1* im Herzgewebe von HLHS-Patienten im Vergleich zu Patienten mit Rechtsherzfehlbildungen herunterreguliert. B: Elektropherogramm der mittels Sanger-Sequenzierung verifizierten Frameshift-Mutation c.592delA (p.S198Afs*8) in *SOSTDC1* bei HLHS_132 (im IAS-Gewebe und in pB detektiert). Dargestellt ist oben die Sequenz des Vorwärts-Primers (VP) und unten die des Rückwärts-Primers (RP). Rote Buchstaben stellen die mutierte Base dar, unterstrichene Buchstaben die mutierte Kodonsequenz der Aminosäuren. Das durch die Frameshift-Mutation resultierende vorzeitige Stopkodon ist in der VP-Sequenz markiert (TGA). C: Stammbaum der Familie 6 (Tabelle 2) mit Genotyp der Mutation c.592delA in *SOSTDC1*. Eine schwarze Füllung symbolisiert den Indexpatienten mit HLHS, während eine weiße Füllung für nicht betroffene Familienmitglieder steht. Der Vater (I:1) wies eine heterozygote Mutation in *ACE2* auf, die Mutter (I:2) zeigte keine Mutation in *SOSTDC1* und der Indexpatient mit HLHS (II:1) zeigte eine heterozygote Ausprägung der Mutation in *SOSTDC1*.

Zusammenfassend konnten zwei Gene – *ACE2* und *SOSTDC1* – bei Patienten mit HLHS mit einer differentiellen Expression im Vergleich zu Rechtsherzfehlbildungen und proteinverändernden Genvarianten, die nach Polyphen-2 als *probably damaging* eingestuft sind, identifiziert werden. Zusätzlich konnte in *SOSTDC1* eine kardiale Dysfunktion im Zebrafisch-Modell (noch nicht publiziert) nachgewiesen werden.

3.3.3 Identifizierung von *XYLT1* mit differenzieller DNA-Methylierung und aberranter Genexpression bei Patienten mit HLHS

In der integrativen Analyse dieses Abschnitts wurden die bei CHD-Patienten identifizierten differenziell methylierten Gene aus T-Test-Analysen des Abschnitts 3.2.2 mit Genen, die eine differenzielle Genexpression im Herzgewebe von HLHS-Patienten im Vergleich zu Patienten mit Rechtsherzfehlbildungen zeigten (RNA-Sequenzierung, siehe 3.2.3), abgeglichen. Das Gen *XYLT1* (NM_022166) zeigte hierbei eine signifikante Hypomethylierung zweier DM-CpG-Loci im Herzgewebe von HLHS-Patienten im Vergleich zu fetalem Herzkontrollgewebe (siehe 3.2.2.5, Boxplot in Abbildung 46 und Verifizierung in Abbildung 47). Zusätzlich konnte in demselben analysierten IAS-Gewebe der HLHS-Patienten eine signifikante Hochregulation der Genexpression ($p = 0,002865$) im Vergleich zu Patienten mit Rechtsherzfehlbildungen detektiert werden (Abbildung 56).

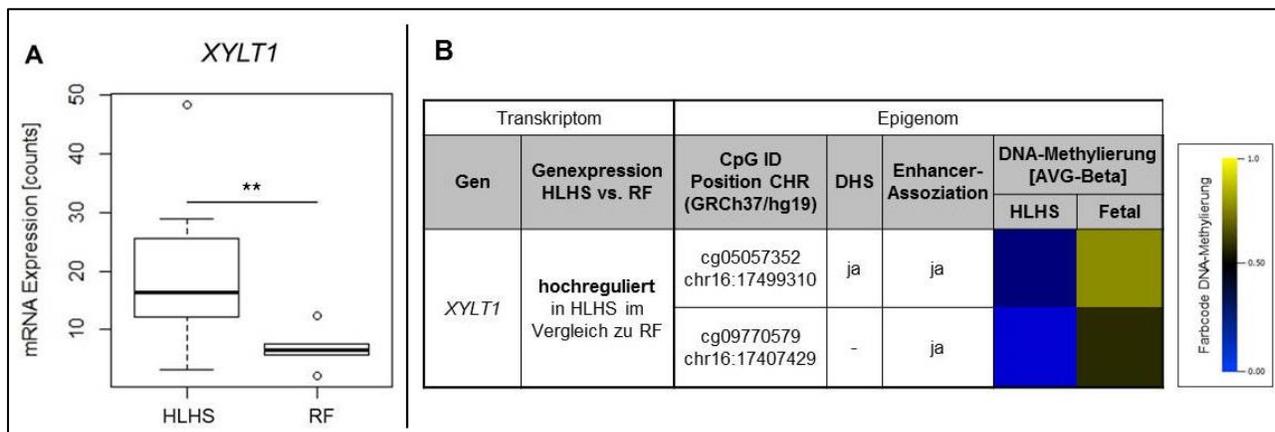


Abbildung 56 (A) Boxplot-Darstellung der differentiellen Genexpression von *XYLT1* im IAS-Gewebe von HLHS-Patienten im Vergleich zu Patienten mit Rechtsherzfehlbildungen sowie (B) differenzielle Methylierung in *XYLT1*.

A: Die mRNA Expression der RNA-Sequenzierungsanalyse (siehe 3.2.3), ist in *counts* (Anzahl gemappter *Reads*) auf der y-Achse angegeben. Im *Fisher's Exact* Test lag eine signifikant hochregulierte Expression von *XYLT1* bei HLHS-Patienten ($N = 26$) im Vergleich zu Patienten mit Rechtsherzfehlbildungen (RF) ($N = 10$) vor ($p = 0,002865$). B: DM-CpG-Loci von *XYLT1* (aus HLHS-Fetal-Analyse siehe 3.2.2.5). Aufgeführt sind die CpG-Loci in *XYLT1* und deren chromosomale Position (CpG ID Position CHR) sowie Angaben zu DHS (*DNase I hypersensitive site*), und Enhancer-Assoziation nach Illumina-Annotation pro CpG-Locus. Die DNA-Methylierung (AVG-Beta-Mittelwerte) der Gruppen HLHS und fetales Herzkontrollgewebe (Fetal) ist entsprechend der Farbcode-Legende rechts der Abbildung eingefärbt.

Zusammenfassend konnte im IAS-Gewebe von HLHS-Patienten in *XYLT1* sowohl eine differenzielle Methylierung von CpG-Loci regulativer DNA-Regionen als auch eine differenzielle Genexpression im Vergleich zu Patienten mit Rechtsherzfehlbildungen detektiert werden.

3.3.4 Gesamtintegrative Analyse

Zur Analyse ursächlicher Zusammenhänge in der Entstehung des HLHS wurde in dieser Arbeit auf verschiedenen Ebenen der Genregulation, d.h. auf Genom-, Epigenom- und Transkriptomebene, Analysen des Herzgewebes von Patienten mit HLHS durchgeführt. Zwar lagen jeweils Überlappungen von potenziell krankheitsrelevanten Genen vor, die in genetischen und epigenetischen Analysen (siehe 3.3.1), in genetischen und transkriptionellen Analysen (siehe 3.3.2) sowie in epigenetischen und transkriptionellen Analysen (siehe 3.3.3) identifiziert wurden, es konnte jedoch kein Gen identifiziert werden, welches auf allen drei Ebenen Veränderungen aufwies. Konvergierende Gene mit kardialem Phänotyp im Mausmodell bzw. mit aus der Literatur bekannten Assoziationen zu CHDs (Herz-Genliste) konnten dagegen auf allen drei Ebenen festgestellt werden. Betrachtete man betroffene Gennetzwerke und *Pathways*, so waren vornehmlich in Entwicklungsprozessen involvierte Gene angereichert, bezüglich der kardiovaskulären Entwicklung speziell auf Epigenomebene (Abbildung 57).

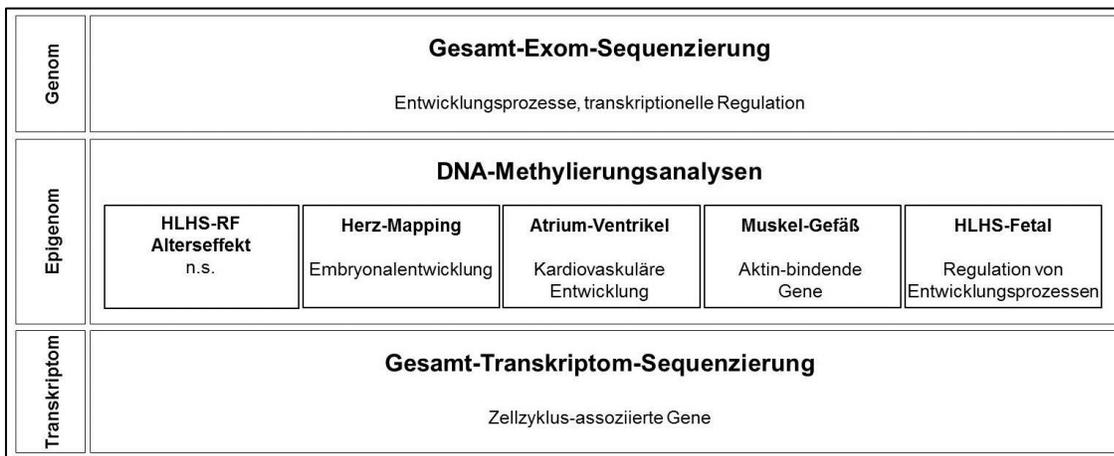


Abbildung 57 Übersicht der Gennetzwerke (bei Genom- und Transkriptom-Analysen) bzw. *Pathways* (bei Epigenom-Analysen), die in Genen angereichert sind, welche in den verschiedenen Analysen auf Genom-, Epigenom- und Transkriptomebene identifiziert wurden. Teilprojektbezeichnungen der DNA-Methylierungsanalysen entsprechend Abschnitt 3.2.2 (Tabelle 26); n.s.: Anreicherungsanalyse nicht signifikant

Zusammenfassend konnten in dieser Arbeit auf Keimbahn-Ebene potenziell krankheitsrelevante und teilweise kausale Genvarianten bei Fällen mit syndrom-assoziierten und nicht-syndromalen CHDs nachgewiesen werden. Darüber hinaus konnten auf somatischer Ebene differenzielle DNA-Methylierungsprofile bei Patienten mit HLHS und verschiedenen CHDs, funktionelle Genvarianten sowie eine aberrante Genexpression bei Patienten mit HLHS detektiert werden. Es konnten Überlappungen identifizierter potenziell krankheitsrelevanter Gene auf Genom-, Epigenom- und Transkriptomebene festgestellt werden.

4 Diskussion

In der vorliegenden Arbeit wurden Analysen auf Genom-, Epigenom- und Transkriptomebene an Herzgewebe- bzw. Keimbahnmaterial durchgeführt, um die molekularen Mechanismen in der Entstehung des HLHS zu untersuchen, dessen Ätiologie bis heute weitgehend unklar ist [188]. Es konnten in dieser Arbeit auf Keimbahn-Ebene kausale Genvarianten bei Fällen eines syndrom-assoziierten CHDs und bei nicht-syndromalen HLHS-Fällen identifiziert werden. Darüber hinaus ließen sich auf somatischer Ebene im Herzgewebe von Patienten mit HLHS potenziell krankheitsrelevante Genvarianten sowie eine differenzielle DNA-Methylierung und Genexpression detektieren. Des Weiteren konnten differenzielle DNA-Methylierungsprofile in unterschiedlichen Regionen des Herzens bei einer Analyse von Herzgewebeproben von Patienten mit verschiedenen CHDs nachgewiesen werden. In einer integrativen Analyse ließen sich Überlappungen identifizierter Gene auf Genom-, Epigenom- und Transkriptomebene feststellen. Zusammengefasst konnten folgende Ergebnisse erzielt werden:

- (1) In der vorliegenden Arbeit konnten kausale Genvarianten in *TGDS* im Rahmen der Studie von Ehmke et al. (2014) [104] nach autosomal-rezessivem Vererbungsmodell in drei Familien mit Catel-Manzke-Syndrom als Beispiel eines syndrom-assoziierten CHDs identifiziert und verifiziert werden.
- (2) In *NR2F2* ließ sich eine bisher unbekannte kausale *de novo* Splice-Donor-Mutation bei einem Patienten mit HLHS als Beispiel eines nicht-syndromalen CHDs als Beitrag zur Studie von Al Turki et al. (2014) [303] nachweisen, welche einen ursächlichen Zusammenhang von *NR2F2*-Genvarianten mit isoliert auftretenden AVSD- und LVOTO-Fällen zeigen konnte.
- (3) Im Rahmen einer WES-Studie zu Familien mit LVOTOs konnten bei einer Familie mit zwei HLHS-Fällen potenziell krankheitsrelevante Genvarianten in *C15orf62*, *TNKS1BP1* und *PRF1* nach autosomal dominantem Vererbungsmodell mit reduzierter Penetranz identifiziert und verifiziert werden.
- (4) Im Herzgewebe von Patienten mit HLHS konnten mittels WES-Analyse insgesamt 685 Gene mit seltenen bzw. bisher unbekanntem potenziell krankheitsrelevanten Mutationen detektiert werden. Hierbei ließen sich Genvarianten in Genen, die in Entwicklungsprozessen involviert sind, wie Notch- und Wnt-Signalwege, sowie in Genen der transkriptionellen und posttranskriptionellen Regulation identifizieren. 17/685 Gene waren mit einem

kardiovaskulären Phänotyp im Mausmodell nach MGI-Datenbank assoziiert und 10/685 Gene wurden zuvor bei Fällen mit kardiovaskulären Erkrankungen (nach Literaturrecherche NCBI PubMed-Datenbank) beschrieben.

- (5) Mithilfe der WES-Methode detektierte potenziell somatische Genvarianten ließen auf das mögliche Vorhandensein von Herzgewebe-spezifischen Mutationen in *ANKRD11* und *BHLHA15* schließen. Jedoch war eine Detektion dieser Mutationen mittels Sanger-Sequenzierungs-Technik zum bisherigen Zeitpunkt nicht möglich. Deren Nachweis erfordert zukünftig *Deep-Sequencing*-Technologien.
- (6) Der Vergleich der DNA-Methylierung im Herzgewebe von HLHS-Patienten und Patienten mit Rechtsherzfehlbildungen ergab keine signifikante differenzielle Methylierung, jedoch ließ sich ein altersabhängiges Methylierungsprofil im Herzgewebe identifizieren.
- (7) Es konnte erstmalig an primären Herzgewebeproben gezeigt werden, dass unterschiedliche Herzregionen gewebespezifisch ein differenzielles DNA-Methylierungsprofil aufweisen. Hierbei ließen sich zum einen Atrium- und Ventrikelspezifische Methylierungsmuster, und zum anderen eine differenzielle Muskel- und Gefäßsignatur detektieren.
- (8) Im Herzgewebe von Patienten mit HLHS verglichen mit fetalem Herzkontrollgewebe konnten multiple differenziell methylierte CpG-Loci in den Genen *NR2F1*, *NR2F1-AS1*, *PDXDC1* und *XYLT1* identifiziert werden. Insgesamt ließen sich hohe Methylierungsdifferenzen von $\Delta\beta > 50\%$ im Herzgewebe von HLHS-Patienten verglichen mit fetalem Herzkontrollgewebe detektieren.
- (9) In einem möglichen Zusammenhang mit der Entstehung des HLHS konnten verschiedene Gene der Regulation der Transkription (*NR2F2* und *ACE2*), der Ausbildung kardialer Strukturen (*DOCK1* und *XYLT1*) sowie der Signaltransduktion (*NOTCH1*, *CTNNB1* und *SOSTDC1*) in integrativen Analysen auf Genom-, Epigenom- oder Transkriptomebene im Herzgewebe von HLHS-Patienten identifiziert werden.

In diesem Teil der Arbeit werden die Ergebnisse im Kontext bisheriger wissenschaftlicher Erkenntnisse diskutiert. Hierbei werden zunächst die methodischen Aspekte der angewandten Methoden erörtert und anschließend auf die biologische Bedeutung der in dieser Arbeit gewonnenen Erkenntnisse eingegangen.

4.1 Methodische Aspekte

Im folgenden Abschnitt werden die methodischen und technischen Aspekte dieser Arbeit diskutiert. Hierbei wird zum einen auf die Herzgewebeprobenentnahme und das Probenmaterial und zum anderen auf die Labortechniken von *Next Generation Sequencing*-Technologien, wie der Gesamt-Exom-Sequenzierung und der RNA-Sequenzierung, sowie auf Methoden der DNA-Methylierungsanalyse, d.h. der Natriumbisulfit-Konversion, des 450K-Arrays, der Bisulfit-Pyrosequenzierung (BSPS) und der Epityper-Technik, eingegangen.

4.1.1 Gewebeisolation und Herzgewebetyp

In der vorliegenden Arbeit wurden Herzgewebeproben von HLHS-Patienten sowie anderer CHD-Typen analysiert, um die pathogenen molekularen Mechanismen im betroffenen Gewebe zu untersuchen. Hierbei wurde das interatriale Septum (IAS) als Analysegewebe zur Untersuchung des HLHS verwendet. Die Morphologie des IAS stellt einen wichtigen Faktor in der Pathogenese dieser Herzfehlbildung dar, indem eine variable Morphologie des IAS in einem möglichen Zusammenhang mit einer aberranten kardialen Entwicklung detektiert werden konnte [336]. Sowohl anatomische Anomalien des IAS [198], als auch eine verringerte Größe des *Foramen ovale* [197] konnten in Assoziation mit dem HLHS beobachtet werden. Eine daraus resultierende Beeinträchtigung des fetalen Rechts-Links-Shunts könnte eine linksventrikuläre Wachstumsinderung zur Folge haben [336]. Diese genannten Aspekte sowie die regelmäßige Verfügbarkeit des Gewebes aufgrund der Atrioseptektomie im Rahmen des Operationsverfahrens (siehe Abschnitt 8 Tabelle 30) begründen die Auswahl des IAS als Analysegewebe. In Anbetracht der Hypothese, dass „primäre“ Defekte in Ein- und Ausflusstrakt des LVs konsekutiv sekundäre Veränderungen der Flussdynamik und so eine abnorme Entwicklung linksseitiger Strukturen bewirken [200], wäre vermutlich eine Analyse letztgenannter anatomischer Strukturen von größerer Aussagekraft. Diese können jedoch im Rahmen der operativen Behandlung des HLHS nicht gewonnen werden, sodass ihre Analyse lediglich mit *post mortem* Proben oder im Rahmen seltener Herztransplantationen gewonnenen Proben möglich wäre. Jedoch lagen solche Proben von Patienten mit HLHS für diese Arbeit nicht vor.

Das Patientenalter bei Probenentnahme reicht von ≤ 3 Wochen bei 69 % (20/29) der HLHS-Fälle bis zu > 30 Wochen bei 10 % (3/29) der HLHS-Fälle. Da in der frühen neonatalen Phase wesentliche Veränderungen des Genexpressionsprofils des Herzens [337 - 339], sowie

weitreichende dynamische Veränderungen des kardialen Methyloms stattfinden [319], stellt die Heterogenität des Gewebeaters einen Einflussfaktor auf Analysen des Epigenoms und Transkriptoms dar. Solche Alterseffekte waren in dieser Arbeit in DNA-Methylierungsanalysen von IAS-Gewebeproben von HLHS-Patienten und Patienten mit Rechtsherzfehlbildungen (siehe 3.2.2.1) detektierbar, deren Ergebnisse in Abschnitt 4.2.3.2 diskutiert werden. Grundsätzlich wäre ein einheitliches Gewebealter erstrebenswert gewesen, jedoch handelt es sich beim HLHS mit einer Prävalenz von 2 – 3 Fällen pro 10.000 Lebendgeburten [183, 186, 187] um eine seltene Herzfehlbildung, sodass nur wenige Gewebeproben zur Verfügung stehen. Zudem sind die Gewebeproben der Atrioseptektomie mehrheitlich kleiner als Reiskorn-Größe, sodass ein Großteil der Proben für weitere gesamtgenomische Analysen nicht eingesetzt werden konnten.

Einen weiteren Einflussfaktor stellt die Heterogenität der Herzgewebeproben in Anbetracht des Zelltyps dar. IAS-Gewebeproben bestehen hauptsächlich aus einem muskulären Anteil mit Kardiomyozyten, sowie aus Endothel und Fibroblasten [248], deren Anteil pro Gewebeprobe leicht variieren kann. Aufgrund des limitierten Probenmaterials war eine detailliertere Analyse der zellulären Zusammensetzung nicht möglich, sodass der variable Anteil unterschiedlicher Zelltypen potenziell einen Einflussfaktor darstellt, der in dieser Arbeit durchgeführten Analysen darstellen könnte.

Als Kontrollkohorte wurden für die initiale Methylierungsanalyse der HLHS-Fälle Herzgewebeproben von Patienten mit Fehlbildungen der rechten Herzhälfte verwendet. Zwar wurde Gewebe derselben Herzregion (IAS) in beiden Kohorten analysiert, sodass gewebespezifische Effekte ausgeschlossen werden konnten, jedoch handelte es sich bei der Kontroll-Kohorte um herzkranken Individuen, sodass kein Vergleich von „krankem“ mit „gesundem“ Gewebe möglich war. Das Akquirieren von Herzgewebeproben gesunder Individuen ist naturgemäß nicht möglich, da keine Indikation für eine Herzoperation besteht. Zwar kann *post mortem* Gewebe herzgesunder Individuen gewonnen werden, jedoch liegt dies von Neugeborenen nur sehr selten vor und solche gesunde Herzen werden, wenn möglich, für Transplantationszwecke verwendet. Bei Analyse solcher schwer zugänglichen, seltenen primären Gewebeproben ist somit grundsätzlich eine Limitierung des experimentellen Designs und der Auswertung gegeben, da kein Vergleich zu Kontrollen im herkömmlichen Sinne möglich ist.

Post mortem Herzgewebeproben eines DCM-Patienten wurden in dieser Arbeit zur Analyse der Herzregion-Methylierung verwendet. In der Studie von Rhein et al. (2015) [340] konnte gezeigt werden, dass die Zeitspanne nach Todeseintritt (*post mortem interval* PMI) potenziell einen Einfluss auf die gemessene DNA-Methylierung im Gewebe (hier Gehirngewebe von Schweinen)

hat. Eine hohe Integrität der gemessenen DNA-Methylierung mittels Bisulfit-Pyrosequenzierung konnte bis zu einem Zeitpunkt von 72 h PMI gezeigt werden, während die Varianz der Messwerte mit höherem PMI deutlich zunahm [340]. Auch in der Studie von Barrachina and Ferrer (2009) [341] zu Patienten mit Alzheimer-Erkrankung konnte eine hohe Reproduzierbarkeit der gemessenen DNA-Methylierung in Gehirnproben, die bis zu einem Zeitpunkt von 48 h PMI entnommen wurden, gezeigt werden. Die in dieser Arbeit verwendeten *post mortem* Herzgewebeproben des DCM-Patienten wurden zu einem Zeitpunkt von 12 h PMI entnommen. Somit konnte davon ausgegangen werden, dass die *post mortem* Exzision keinen signifikanten Effekt auf die Ergebnisse der durchgeführten DNA-Methylierungsanalysen hatte.

4.1.2 Next Generation Sequencing-Technologien

Das *Next Generation Sequencing* (NGS), erstmals im Jahr 2005 als sogenannte zyklische Array-Strategie beschrieben [342, 343], stellt als nachfolgende Technologie der „*first generation*“ Sequenzier-Methode, der Sanger-Sequenzierung [344], eine Hochdurchsatz-Technologie dar, die eine massive parallele Sequenzierung von Millionen von DNA-Fragmenten in einem einzigen Sequenzierlauf ermöglicht. Die NGS-Technologie umfasst Fortschritte der Sequenzier-Methodik, Signal-Detektion und Rechenleistung, sodass eine klonale Sequenzierung mit einer erhöhten Sensitivität durch parallele Sequenzierung die Generierung großer Datenmengen DNA- oder RNA-basierter Sequenzen ermöglicht [345]. Die drei Arbeitsabläufe der NGS-Technologie – *Library*-Präparation, Sequenzierung und bioinformatische Auswertung – umfassen kritische Schritte, die die Ergebnisse und deren Auswertung beeinflussen. Für die beiden NGS-Technologien, der Gesamt-Exom-Sequenzierung und der RNA-Sequenzierung, die in dieser Arbeit angewendet wurden, wird dies im Folgenden diskutiert.

4.1.2.1 Gesamt-Exom-Sequenzierungstechnologie

Die Gesamt-Exom-Sequenzierung (WES) ist eine Technik zur Sequenzierung aller bekannten protein-kodierenden Gene eines Genoms. Lediglich 2,94 % des Genoms sind kodierend für exonische Sequenzen (ca. 25.000 Gene), bzw. 1,2 % des Genoms bei ausschließlicher Betrachtung der Aminosäure-kodierenden DNA-Sequenz, d.h. ohne UTR-Regionen [346]. Zwar stellen diese Sequenzen nur einen kleinen Anteil des Genoms dar, Mutationen in diesen sind jedoch mit höherer Wahrscheinlichkeit von schwerwiegender Konsequenz als in Sequenzen der

restlichen 98 % des Genoms [347, 348]. Ungefähr 80 – 90 % der bekannten krankheitsrelevanten Mutationen liegen innerhalb exonischer Sequenzen und mithilfe der WES-Technologie ließen sich seit 2009 über 30 neue kausale Gene Mendelscher Erkrankungen identifizieren [349, 350]. Die WES-Technologie stellt insgesamt eine kosteneffektive Methode zur Identifizierung klinisch relevanter Genvarianten dar. Um einen validen Datensatz zu erhalten, müssen jedoch einige Faktoren in den Arbeitsabläufen der WES-Analyse beachtet werden.

Limitierungen der *Library*-Präparation

Die *Library*-Präparation (in Abschnitt 2.2.1.7 beschrieben), welche eine Anreicherung exonischer Sequenzen sowie die Adapterligation für die jeweilige Sequenzier-Plattform umfasst, stellt eine wesentliche Komponente in der WES-Analyse dar. Unterschiedliche *Library*-Präparationen, die auf verschiedenen Enrichment-Kits basieren, zeigten in dieser Arbeit teilweise deutliche Unterschiede in der *Reads Coverage* bei gleicher Sequenzier-Plattform, hier dem Illumina HiSeq 2000-System. Vier von 11 analysierten Genvarianten, die zwar über die Plattform des „NimbleGen“-Kits detektierbar waren, wiesen keine oder nur eine geringe *Coverage* bei der „Agilent“-Plattform auf (siehe 3.2.1.2). Vergleicht man die beiden Anreicherungs-Plattformen, so unterscheiden sich diese in Zielsequenz-Regionen, der *Bait*-Länge und –Dichte, sowie in der *Capture*-Art [351 - 353]. Bei der Agilent-Technologie werden cRNA-*Baits*, d.h. ein Oligo-Pool aus cRNA (RNA erzeugt aus cDNA) komplementär zu den exonischen Ziel-Sequenzen, und bei der NimbleGen-Technologie DNA-*Baits* zur Anreicherung der Ziel-Sequenzen verwendet. Die NimbleGen-Plattform beinhaltet überlappende *Baits* in hoher Dichte (1,2 Mio. *Baits*), während die Agilent-Plattform keine überlappenden *Baits* (650.000 *Baits*) enthält. Zwar deckt die NimbleGen-Plattform weniger genomische Regionen ab (~44 Mb) als die Agilent-Plattform (51 Mb), jedoch lassen sich SNPs mit einer höheren Sensitivität und geringerer Sequenzierhäufigkeit mithilfe der NimbleGen-Plattform detektieren [351]. Die beiden Anreicherungs-Plattformen unterscheiden sich zudem in den Zielsequenz-Regionen, da jede Plattform diese Sequenz-Regionen basierend auf einer unterschiedlichen Kombination aus den verfügbaren Datenbanken zu mRNA-kodierenden Sequenzen – inklusive UCSC KnownGenes [354], RefSeq [355] und Ensembl [356] – definiert. Die beiden Plattformen zeigen eine Überlappung von ~39 Mb, sodass die restlichen genomischen Regionen jeweils nur von einer der beiden Plattformen abgedeckt werden [351]. Somit ist die Vergleichbarkeit der Sequenzier-Daten zwischen verschiedenen Plattformen teilweise limitiert und muss in den jeweiligen Analysen in Betracht gezogen werden.

Möglichkeiten und Limitierungen bei der Analyse somatischer Mutationen

Eine kombinierte Keimbahn- und somatische Analyse einer Familie mit HLHS wurde durchgeführt, um herzwegewebe-spezifische somatische Mutationen zu identifizieren (siehe 3.2.1.2). Dabei konnten die zuvor mittels WES-Technologie im IAS-Gewebe detektierten Mutationen anschließend mithilfe der Methode der Sanger-Sequenzierung nicht repliziert werden. Dies ließe sich dadurch erklären, dass die Sanger-Sequenzierungs-Technik eine geringere Sensitivität als die NGS-Technik der WES-Analyse aufweist. Die Sanger-Sequenzierung hat eine Sensitivität von minimal 10 – 20 % mutierter AF, während derzeitige *Deep-Sequencing* NGS-Methoden eine Detektion von somatischen Mosaik-Mutationen mit einer AF von 1 % ermöglichen [357]. Beispielsweise ließ sich in einer Studie von Miyatake et al. (2014) [358] eine neue krankheitsassoziierte Mutation in *ACTA1* bei zwei Geschwistern mit Nermalin-Myopathie, einer kongenitalen Myopathie, nachweisen, während mithilfe der angewandten Technik der Sanger-Sequenzierung kein Nachweis der Mutation bei den Eltern möglich war. Jedoch konnten die Autoren anschließend durch Anwendung der *Deep-Sequencing*-Technologie ein somatisches Mosaik dieser Mutation in *ACTA1* bei der phänotypisch nicht-betroffenen Mutter nachweisen, sodass solche somatischen Mutationen bei höherer Sequenzierentiefe detektierbar sind. Somatische Mutationen in Hautgewebe ließen sich kürzlich mit einer *Coverage* von 500x nachweisen [359]. Zur gezielten Analyse somatischer Mutationen in Herzgewebeproben von CHD-Patienten sind daher wahrscheinlich minimale Sequenzierertiefen von 500x nötig, um potenziell krankheitsrelevante somatische Mutationen nachzuweisen.

Herausforderungen in der Interpretation von Sequenzierdaten

Die Analyse großer Sequenzier-Datenmengen der NGS-Technologie stellt eine wesentliche Herausforderung in der Interpretation dieser Daten dar und hängt entscheidend von einem akkuraten *Varianten-Calling* ab [360]. Insbesondere bei einer *Coverage* ≤ 6 besteht eine hohe Wahrscheinlichkeit, dass nur eins der beiden Allele des diploiden Individuums in der analysierten Region der *Reads* vorliegt, was eine Schwierigkeit im *Varianten-Calling* darstellt. Um ein valides Ergebnis zu gewährleisten, wurde in dieser Arbeit im Rahmen der Varianten-Filterung (siehe 2.2.1.7) eine *Calling*-Software-spezifische Selektion von SNPs und INDELS durchgeführt, mit der Detektion durch minimal zwei bzw. drei *Calling*-Software-Tools pro Variante. GATK [361, 362] zeigte in verschiedenen Analysen von Liu et al. (2013a) [363] die höchste Spezifität und Sensitivität im Vergleich zu anderen *Calling*-Software-Tools,

insbesondere bei Datensätzen mit geringer *Coverage* [363 - 365]. Ein *Calling* durch GATK galt daher in dieser Arbeit als Minimum zur Filterung der Varianten. Zusätzlich wurde Mpileup [366, 367] in Kombination mit GATK als SNP-*Calling*-Filter verwendet. Das INDEL-*Calling* zeigt meist eine geringe Konkordanz zwischen verschiedenen *Calling*-Software-Tools [368, 369], was sich durch eine hohe Komplexität und Missalignments in schwierigen genomischen Regionen (z.B. repetitive Sequenzen) erklären lässt. Für INDELS galt somit in dieser Arbeit eine minimale Kombination aus vier verschiedenen *Calling*-Software-Tools, um eine verlässliche Detektion der Genvarianten zu gewährleisten. Hierzu zählten GATK, Mpileup, sowie zusätzliche *Calling*-Software-Tools für INDELS wie Dindel [370] und SplazerS/SNPstore [371].

4.1.2.2 RNA-Sequenzierung

Seit Etablierung der NGS-Technologien ermöglicht die RNA-Sequenzierung, welche auf einer Hochdurchsatz-Sequenzierung von cDNA-*Libraries* basiert, einen großen Fortschritt in der Charakterisierung und Quantifizierung von Transkriptomen [292]. Insbesondere bei dieser NGS-Technologie ist die Qualität der *Library*-Präparation von wesentlicher Bedeutung. Der zu analysierende Pool von Sequenzen beinhaltet eine komplexe Mischung einer unterschiedlichen Anzahl von mRNA-Molekülen, welche die Biologie der differentiellen Genexpression reflektieren [292]. Die Schwierigkeit der *Library*-Präparation besteht darin, diese Komplexität zu bewahren und amplifikationsbasierte klonale Effekte zu minimieren [372]. Auch die cDNA-Synthese stellt einen kritischen Schritt in der RNA-Sequenzierung dar, indem ein sogenanntes *Template Switching* stattfinden kann [373 - 375]. Dies beschreibt einen Prozess während der reversen Transkription, bei dem die sich bildende cDNA von der *Template*-RNA dissoziiert und an einen anderen RNA-Strang mit ähnlicher Sequenz hybridisiert. Auf diese Weise können chimäre cDNAs als Artefakte entstehen und die Identifizierung von Exon-Intron-Übergängen oder „echter“ chimärer Transkripte beeinträchtigen [373]. Ungleichmäßigkeiten in der *Coverage* der Signalintensität über die Transkriptlänge hinweg können neben einer fehlerhaften cDNA-Synthese auch durch ein fehlerhaftes Hexamer-*Priming* (zur cDNA-Synthese) [376, 377], als auch durch Fehler in der Amplifikation der *Library* [372] und in der Sequenzierung [378, 379] hervorgerufen werden. Eine akkurate nachfolgende bioinformatische Auswertung ist daher von wesentlicher Bedeutung für die Analysierbarkeit der Ergebnisse.

Herausforderungen in der Analyse der differentiellen Genexpression

Die Identifizierung differentiell exprimierter Gene, welche eine Normalisierung der gemappten *Reads* pro Transkript basierend auf dessen Länge implementiert, stellt eine wesentliche Herausforderung in der Analyse der RNA-Sequenzierung dar. Verschiedene Software-Anwendungen wurden dazu entwickelt, wie Anwendungen mittels linearer Modellierung, nicht-parametrische Anwendungen sowie Anwendungen basierend auf negativer binomialer Regression (Modellierung von *counts*-Variablen) [380], wie u.a. EdgeR [254] oder Cuffdiff [381], bzw. Cuffdiff2 [382]. Eine typische Expressions-Analyse-Pipeline umfasst nach Trapnell et al. (2012) [383] eine Transkript-Assemblierung mittels Cufflinks gefolgt von einer Annotations-Analyse mittels Cuffmerge und eine Expressions-Analyse durch Cuffdiff(2). Jedoch resultierte die Anwendung dieses Analyse-Ansatzes in einer ungewöhnlich geringen Anzahl differentiell exprimierter Gene in der initialen Analyse der IAS-Gewebeproben. Der Grund dafür kann zum einen in der Variabilität der *Read Counts* und zum anderen darin gelegen haben, dass Cufflinks eine minimale *Coverage* pro Transkript benötigt, sodass es zum Ausschluss des Gens aus der Analyse kommt, sobald nur ein Replikat unterhalb dieses Grenzwerts fällt. Daher wurde in dieser Arbeit die EdgeR-Plattform zur Genexpressions-Analyse verwendet. Zwischen beiden Plattformen liegt eine hohe Konkordanz ($\geq 0,99$ Korrelation) vor [384], in Bezug auf differentielle Expressions-Analysen zeigt EdgeR jedoch Vorteile gegenüber CuffDiff2, da EdgeR eine höhere Toleranz gegenüber einer unbalanzierten *Library*-Größe und geringer Sequenziertiefe (< 10 Mio. *Reads* auf Gen-Level) sowie eine bessere Detektion richtig-positiver differentiell exprimierter Gene aufweist [385].

Mithilfe der RNA-Sequenzierung erfolgt die quantitative Bestimmung der Gesamtheit aller Transkripte zu einem bestimmten Entwicklungsstadium und physiologischen Bedingungen [292], sodass ein einheitliches Alter des zu analysierenden Gewebes, hier des Herzwebes, von wesentlicher Bedeutung ist. Hierbei ist zu beachten, dass in der frühen postnatalen Entwicklung des Herzens dynamische Veränderungen des Genexpressionslevels auftreten [337 – 339, 386], sodass das Herzwewebe zu verschiedenen Zeitpunkten potenziell unterschiedliche Genexpressionsmuster zeigen kann. Diese dynamischen Veränderungen könnten somit eine Ursache für die in dieser Arbeit detektierte Heterogenität der *Ratio*-Werts der mRNA-Expression innerhalb der Patienten-Kohorte der Rechtsherzfehlbildungen, die eine Altersspanne von vier Tagen bis 19 Jahren aufweist, darstellen. Zudem sollten biologische Replikate zur Erhöhung der statistischen Power von Analysen der differentiellen Genexpression eingesetzt werden [387], um einen verlässlichen Datensatz zu gewährleisten. Aufgrund des limitierten Probenmaterials solcher primärer Herzwegewebeproben war jedoch die Anzahl von Proben mit

ausreichend Gewebematerial mit einheitlichem Alter bei Gewebeentnahme begrenzt, sodass auch Proben von Patienten mit höherem Alter in der vorliegenden Arbeit in die Analyse eingeschlossen wurden.

4.1.3 Techniken der DNA-Methylierungs-Analyse

In dieser Arbeit wurden verschiedene Techniken zur DNA-Methylierungsanalyse angewendet. Diese umfassten eine globale Analyse mittels 450K-Array und Locus-spezifische Analysen mittels BSPS oder Epityper-Technik. Die Techniken werden im Folgenden hinsichtlich Vorteilen und Limitierungen diskutiert.

4.1.3.1 Natriumbisulfit-Konversion

Die Natriumbisulfit-Konversion ist Voraussetzung für nachfolgende DNA-Methylierungsanalysen, wie 450K-Array, BSPS und der Epityper-Technik, welche in dieser Arbeit angewendet wurden. Die Natriumbisulfit-Behandlung (siehe 2.2.2.1) bewirkt eine selektive Konversion von Cytosin - nicht jedoch von 5-Methyl-Cytosin – zu Uracil [284, 285], und mittels einer anschließenden PCR-Amplifikation werden Uracil-Moleküle durch Thymin ersetzt. Die Bestimmung der DNA-Methylierung erfolgt anhand des Verhältnisses C/T (bzw. A/G), sodass eine vollständige und fehlerfreie Natriumbisulfit-Konversion für eine valide Analyse der DNA-Methylierung von wesentlicher Bedeutung ist.

Eine unvollständige Natriumbisulfit-Konversion kann beispielsweise durch eine unvollständige Denaturierung von dsDNA, deren Denaturierung Voraussetzung für eine akkurate Konversion von Cytosin zu Uracil ist, oder auch durch Renaturierungsprozesse der DNA hervorgerufen werden [388, 389]. Daraus folgt, dass unmethylierte Cytosine nicht desaminiert werden und somit falsch-positive Ergebnisse ergeben. In dieser Arbeit wurde das EZ DNA Methylation™ Kit (Zymo Research) verwendet, welches in einer komparativen Studie von Konversions-Kits die höchste Konversions-Effizienz von 99,9 % aufwies [390]. Zudem wurde dieses Kit bei jeder der drei verschiedenen angewendeten Methoden der DNA-Methylierungsanalyse eingesetzt, um Effekte durch Anwendung unterschiedlicher Konversions-Kits, welche nach der Studie von Holmes et al. (2014) [390] bei Analyse von verschiedenen Parametern teils heterogene Ergebnisse ergaben, zu vermeiden.

Die Natriumbisulfit-Konversion stellt eine Reaktion unter extremen chemischen Bedingungen dar (u.a. niedriger pH-Wert, hohe Inkubationstemperatur), sodass eine DNA-Degradation auftreten kann [388]. Diese kann u.a. durch Depurinierung und Depyrimidierung mit anschließenden Doppelstrangbrüchen durch N-glykosidische Spaltung ausgelöst werden [391, 392]. Darüber hinaus kann eine fehlerhafte Konversion, d.h., wenn methyliertes Cytosin desaminiert wird und somit Thymin entsteht [388, 393], zu einer Fehlinterpretation des 5-Methyl-Cytosins als „unmethyliertes“ Cytosin führen. Beim EZ DNA Methylation™ Kit (Zymo Research) kann diese fehlerhafte Konversion bis zu 2,7 % ausmachen [390]. Dieser Anteil kann durch eine Verkürzung der Inkubationszeit verringert werden, jedoch auf Kosten der Konversions-Effizienz [390].

Um der beschriebenen Problematik zu begegnen, beinhalten die verschiedenen Techniken der DNA-Methylierungsanalyse Konversionskontrollen zur Qualitätskontrolle. Zum Beispiel beinhaltet der 450K-Array u.a. Bisulfit-Konversionskontrollen (BSK-Sonden, Details siehe 2.2.2.2.1), die einen C/T-Polymorphismus bei ineffizienter Bisulfit-Konversion detektieren. Lässt sich beispielsweise eine Hybridisierung der zu analysierenden, bisulfit-behandelten Proben-DNA an eine unkonvertierte BSK-Sonde als Signal detektieren, so liegt eine unvollständige Konversion dieser Probe vor. In dieser Arbeit kam es dadurch zur Exklusion der Probe HLHS_157, welche eine Signalintensität > 2000 FE bei erwartetem Hintergrundsignal (> 200 FE) zeigte (siehe 3.2.2 Abbildung 25). Auch die BSPS (Details siehe 2.2.2.3) und die Epytper-Technik (Details siehe 2.2.2.4) beinhalten Konversionskontrollen, die für alle analysierten Proben dieser Arbeit eine vollständige Natriumbisulfit-Konversion ergaben.

4.1.3.2 450K-Array

Der 450K-Array ermöglicht als Hochdurchsatz-Technologie die parallele Analyse von bis zu 96 Proben hinsichtlich des DNA-Methylierungsstatus von 485.577 Loci mit einer Abdeckung von 99 % der RefSeq-Gene [355], 96 % bekannter CpG-Inseln der UCSC-Datenbank, sowie weiterer Methylierungs-Loci des Genoms [286]. Fast die Hälfte der Sonden (> 41 %) decken intergenische Regionen ab, wie bspw. bioinformatisch prädiktive Enhancer-, DHS-, oder DMR-Regionen [286, 394]. 65 SNP-Loci ermöglichen zudem eine SNP-Genotypisierung zur Qualitätskontrolle, indem Proben desselben Individuums identische AVG-Beta-Werte für diese Loci aufweisen sollten. In der Studie von Bibikova et al. (2011) [286] wurde die Technologie des 450K-Arrays mit einer hohen Reproduzierbarkeit technischer Replikate ($R^2 > 0,99$) beschrieben. Auch in dieser Arbeit ließ sich eine hohe Korrelation von $R^2 > 0,99$ der technischen Duplikate F_FetalHerz und M_FetalHerz verzeichnen (Abbildung 58). Dasselbe galt für die biologischen

Replikate der Proben HLHS-TGA_117 und Lvan_126, die jeweils für die Gewebetypen LV, RV, LA und RA ein $R^2 > 0,99$ aufwiesen (Abbildung 58). Diese biologischen Replikate wurden jeweils aus topographisch unterschiedlichen Gewebebereichen desselben Gewebetyps im Rahmen der Gewebeexzision aus dem explantierten Herzen entnommen. Die hohe Korrelation der biologischen Replikate rechtfertigt somit die Verwendung des AVG-Beta-Mittelwerts der Methylierungsdaten dieser Proben für die durchgeführten statistischen Analysen (siehe 3.2.2.2 und 3.2.2.3).

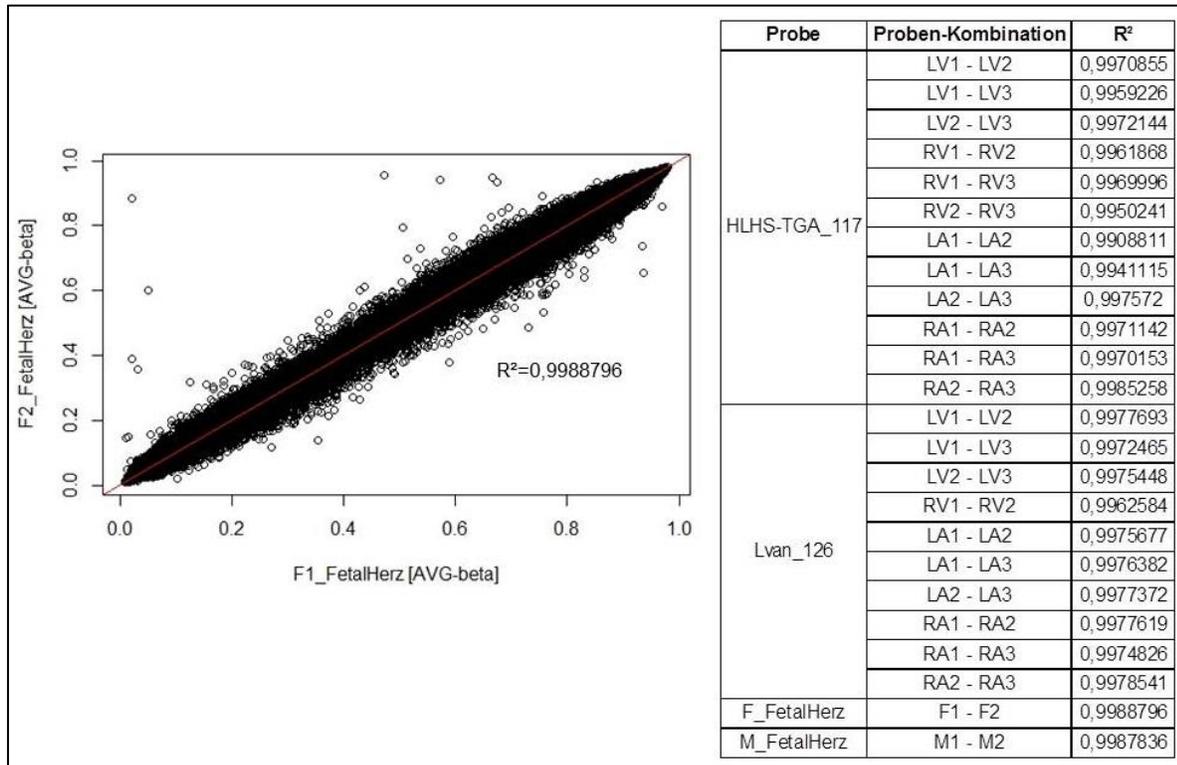


Abbildung 58 Scatterplot der linearen Regression der DNA-Methylierungswerte von 447821 CpG-Loci der technischen und biologischen Replikate der Herz-Mapping-Analyse. Die Methylierungswerte [AVG-Beta] von F1_FetalHerz auf der x-Achse sind gegen die Methylierungswerte von F2_FetalHerz auf der y-Achse zur Qualitätskontrolle der technischen Replikate aus der Herz-Mapping-Analyse (siehe 3.2.2.2) aufgetragen. In der Tabelle rechts des Scatterplots sind die jeweiligen R²-Werte unterschiedlicher Proben-Kombinationen von biologischen Replikaten aus LV-, RV-, LA- und RA-Gewebe der Patientenproben HLHS-TGA_117 und Lvan_126 sowie von technischen Duplikaten der Proben F_FetalHerz und M_Fetalherz dargestellt (berechnet mittels R-Programm). Das Bestimmtheitsmaß R² ergab eine hohe Korrelation mit R² > 0,99 bei technischen und biologischen Replikaten.

Das Sonden-Design des 450K-Arrays beinhaltet eine Präselektion von CpG-Loci, welche zuvor in Methylierungs-basierten Assays identifiziert und durch ein Konsortium epigenetischer Forscher festgelegt wurden [286]. Somit besteht grundsätzlich ein Bias mit nicht-neutraler Hypothese [395]. Im Rahmen der Natrium-Bisulfitkonversion, die einen C/T-Polymorphismus generiert, wird aus einem 4er-Code (A,T,G,C) nahezu ein 3er-Code, indem die restlichen, zuvor methylierten Cytosine nur ca. 3,5 % der totalen Anzahl von Cytosinen des Genoms

ausmachen [396]. Dies erhöht die Wahrscheinlichkeit einer unspezifischen Kohybridisierung der 50mer Oligonukleotidsonden des 450K-Arrays. Ungefähr 8,6 - 25 % der Sonden konnten in verschiedenen Studien als nicht spezifische, d.h. kreuzreaktive Sonden identifiziert werden [397, 398]. Beispielsweise konnte eine bis zu 94 %ige Homologie einiger autosomaler Sondensequenzen zu gonosomalen Sonden detektiert werden [288]. Die DNA-Methylierungswerte solcher kreuzreaktiver Sonden bestehen dabei wahrscheinlich aus einer Kombination der Werte von multiplen Loci, nicht jedoch alleinig der initialen Ziel-CpG-Site [288]. Somit resultieren falsche DNA-Methylierungswerte und artefaktische, differenziell methylierte Loci. Solche kreuzreaktiven Sonden wurden daher in dieser Arbeit entsprechend der Studie von Chen et al. (2013) [288] aus der Analyse ausgeschlossen (Details zu Filterschritten siehe 2.2.2.2.1).

Des Weiteren zeigen ca. 4,3 % der 450K-Sonden eine Überlappung der Ziel-CpG-Site mit bekannten SNPs [288, 397], sodass bei inter-individuellen Analysen artefaktische DNA-Methylierungswerte resultieren. Zudem können SNPs in der Sondensequenz außerhalb der Ziel-CpG-Site auftreten [288, 397]. In dieser Arbeit wurden daher SNPs mit einer AF > 0,05 (1000G rel. 20110521) innerhalb der Distanz von - 3 bp und + 3 bp um die Ziel-CpG-Site aus der Analyse ausgeschlossen (siehe 2.2.2.2.1).

In verschiedenen Studien ließen sich Batch-Effekte bei der 450K-Technologie feststellen, wobei unterschiedliche Experimentatoren, Laborbedingungen, oder Chargen von Reagenzien einen Einfluss haben können [399, 400]. Auch die Array-Position der Probe kann einen Einflussfaktor darstellen [401]. In dieser Arbeit wurden daher möglichst alle Proben pro Teilprojekt gemeinsam hybridisiert und Proben derselben Analysegruppe auf zufällige Positionen des Arrays verteilt. Die Analyse von Proben von unterschiedlichen Arrays und mit unterschiedlichem Hybridisierungsdatum im Rahmen des Herz-Mappings ergaben keine Batch-Effekte (siehe 3.2.2.2 und Abschnitt 8 Abbildung 60).

Ein wichtiger Vorgang in der Prozessierung von 450K-Daten stellt die Normalisierung dar, um technisch-basierte Variationen, wie durch Signaldetektions-Abweichungen und Batch-Effekte, zu vermeiden. Die Infinium I- und Infinium II-Assays [286] (siehe 2.2.2.2) zeigen zudem nach der Studie von Dedeurwaerder et al. (2011) [402] eine unterschiedliche Spanne möglicher DNA-Methylierungswerte, indem der Infinium II-Assay eine geringere Sensitivität in der Detektion extremer Methylierungswerte (AVG-Beta-Werte von 0 bzw. 1) aufweist als der Infinium I-Assay. Somit ist eine Normalisierung von wesentlicher Bedeutung, die in dieser Arbeit mithilfe des RnBeads *R-Package* für jedes einzelne Teilprojekt durchgeführt wurde. Mithilfe der SWAN-

Methode [287] konnten hierbei eine Hintergrundsignal-Korrektur, eine Fluoreszenzsignal-Anpassung und Infinium I/II-Korrektur (Angleichen der Beta-Wert-Verteilung von Infinium I- und Infinium II-Assays) erreicht werden.

Zwar stellt der 450K-Array eine kosteneffektive Hochdurchsatz-Technologie zur globalen DNA-Methylierungsanalyse dar, jedoch deckt dieser Array lediglich ~1,7 % aller CpG-Loci des Genoms ab. Zwar lassen sich nahezu alle RefSeq-Gene [355] und CpG-Inseln, sowie einige genregulatorische Elemente abdecken, zusätzlich könnten jedoch aberrante Methylierungsmuster in den restlichen ~ 98 % des Genoms eine potenzielle krankheitsrelevante Funktion haben [403]. Eine Analyse des Methylierungsstatus des gesamten Genoms mittels Sequenzierungs-basierter Technologien, wie z.B. der Gesamt-Genom-Bisulfit-Sequenzierung, wäre somit in zukünftigen Analysen erstrebenswert, um die ursächlichen Zusammenhänge in der Entstehung komplexer Erkrankungen wie dem HLHS auf gesamtgenomischer Ebene epigenetisch erforschen zu können.

4.1.3.3 Verifizierung von 450K-Ergebnissen mittels Bisulfit-Pyrosequenzierung

Die Bisulfit-Pyrosequenzierung (BSPS), die auf einer *Sequencing-by-Synthesis* Methode basiert [289], ermöglicht die quantitative DNA-Methylierungsanalyse von CpG-Loci und wurde in dieser Arbeit zur Verifizierung von 450K-Ergebnissen verwendet. Eine hohe Übereinstimmung der mit beiden Techniken erhobenen DNA-Methylierungswerte konnte in verschiedenen Analysen der vorliegenden Arbeit nachgewiesen werden. Die differenziell methylierten CpG-Loci im Herzgewebe von HLHS-Patienten verglichen mit fetalem Herzkontrollgewebe der Gene *NR2F1*, *NR2F1-AS1*, *XYLT1*, *PDXDC1*, *CTNNB1* und *STEAP2* zeigten mit einem Bestimmtheitsmaß von $R^2 = 0,94$ bis $R^2 = 0,99$ eine hohe Korrelation der Daten (siehe 3.2.2.5.3 Abbildung 47 bzw. Abschnitt 8 Abbildung 65). Somit ließen sich die DNA-Methylierungswerte des 450K-Arrays gut replizieren und die Ergebnisse dieses Arrays konnten als valide eingestuft werden. Auch in der Literatur lässt sich eine hohe Korrelation der Daten dieser beiden Techniken beim Vergleich der Methylierungswerte desselben CpG-Locus mit $R^2 = 0,88$ verzeichnen [404].

Grundsätzlich gibt es jedoch auch Limitierungen dieser Technik in Bezug auf das Assay-Design. Aufgrund der thermalen Instabilität der Luziferase muss die Pyrosequenzierung bei geringer Temperatur (28 °C) durchgeführt werden und zudem besteht eine Limitierung der Amplifikat-Länge von maximal 350 bp, um die Bildung von Sekundärstrukturen zu vermeiden [290]. Des

Weiteren gestaltet sich das Assay-Design bei GC-reichen Regionen problematisch, sodass in diesen Fällen teilweise die Generierung eines spezifischen PCR-Produkts nicht möglich ist. Dennoch stellt die BSPS insgesamt eine valide Methode zur spezifischen Analyse bestimmter Loci und zur Verifizierung von 450K-Ergebnissen mit einer hohen Korrelation der Daten, wie in dieser Arbeit gezeigt werden konnte, dar.

4.1.3.4 Anwendung der Epityper-Technik zur Verifizierung von 450K-Ergebnissen

Die Epityper-Technik stellt ein massenspektrometrisches Verfahren zur semi-quantitativen DNA-Methylierungsanalyse dar, die auf einer basenspezifischen Spaltung mit anschließender MALDI-TOF-Massenspektrometrie basiert [291]. Der C/T-Polymorphismus lässt sich hier durch eine 16 Da-Massendifferenz messen. Im Rahmen der vorliegenden Arbeit wurde diese Technik neu etabliert und für die Verifizierung von 450K-Array-Ergebnissen angewendet. Mit analysierbaren Amplifikaten von bis zu 600 bp [405] können multiple CpG-Loci parallel in bis zu 384 Proben untersucht werden, sodass diese Technologie eine kosteneffiziente Alternative zur BSPS mit einem höheren Probendurchsatz darstellt. Dennoch lassen sich bei dieser massenspektrometrischen Methode einige Limitierungen feststellen. Zwar beinhaltet das Assay-Design keine zusätzlichen Sequenzierprimer und es entfallen zudem Optimierungs-PCR-Vorgänge (diese sind bei der Epityper-Technik *in silico*-basiert, siehe 2.2.5.1.1), wie dies für die BSPS notwendig ist, jedoch zeigte sich in der Etablierung der Technik im Rahmen dieser Arbeit, dass eine Vielzahl zur Verifizierung ausgewählter CpG-Loci sich aus technischen Gründen nicht verifizieren ließ.

Es ist festzustellen, dass die Analysierbarkeit einiger CpG-Loci eines zu untersuchenden Amplifikats teilweise nicht möglich ist. Zum einen können grundsätzlich Fragmente derselben Masse (*silent peaks*) auftreten und die Detektion des zu analysierenden CpGs im analytischen Fragment durch Reduktion des quantitativen Auflösungsvermögens verfälschen. Zum anderen lassen sich CpGs innerhalb von Fragmenten mit einer zu geringen (< 1500 Da) oder zu großen Masse (> 7000 Da) nicht zuverlässig detektieren [291]. Als Kriterium zur Analysierbarkeit der CpGs galt in dieser Arbeit eine maximale Überlappung von einem Fragment, um artefaktische Messungen zu vermeiden (Details siehe 2.2.2.4 Tabelle 17). Des Weiteren lassen sich CpGs, die sich in räumlicher Nähe zueinander befinden, nur limitiert detektieren, da mit höherer CpG-Dichte pro analytischem Fragment keine valide Detektion des Methylierungsstatus möglich ist. In

dieser Arbeit ließen sich daher nur einige CpG-Loci mithilfe dieser Methode verifizieren, da aus den oben genannten technischen Gründen viele CpG-Loci nicht analysierbar waren.

Im Rahmen der Verifizierung von CpG-Loci in *FGFR2* und *IRX4*, die zwischen atrialem und ventrikulärem Gewebe differenziell methyliert vorlagen, konnte eine hohe Korrelation der Techniken 450K-Array und Epityper mit $R^2 = 0,85$ bzw. $R^2 = 0,94$ in der Regressionsanalyse gezeigt werden (siehe 3.2.2.3.3 Abbildung 39). Jedoch ließ sich eine Streuung der Daten in der linearen Regression feststellen. Zwar blieb die mittels 450K-Array detektierte Methylierungsdifferenz zwischen Atrium- und Ventrikel-Proben von $\Delta\beta$ 0,3 bis 0,4 annähernd gleich, jedoch fielen die Rohwerte bei der Epityper-Technik tendenziell geringer aus (um bis zu $\beta = 0,18$) (siehe Abschnitt 8, Abbildung 61). Dies könnte darin begründet sein, dass das analytische Fragment der analysierten CpG-Loci von *FGFR2* und *IRX4* ein zusätzliches überlappendes Fragment aufwies. Zwar gelten nach „Sequenom Product Preview Note“ von Ehrich et al. (2005) [291] überlappende Signalintensitäten solcher *silent peaks* als auswertbar, jedoch scheint dies in dieser Arbeit einen Einfluss auf die quantitative Detektion gezeigt zu haben. Verschiedene Studien zeigen eine variable CpG-Site-Spezifität der Epityper-Technik, mit einer Konkordanzspanne von 75 % bis 95 % bei Verifizierung mithilfe anderer Methylierungs-Analyse-Methoden [406, 407]. Dies entspricht der in dieser Arbeit detektierten Konkordanz von 82 % bis 99 % zwischen der Epityper-Technik und dem 450K-Array.

Insgesamt stellt die Epityper-Technik eine kosteneffiziente Hochdurchsatz-Technologie mit der Möglichkeit der Analyse von großen Amplifikaten von bis zu 600 bp dar. Zur Verifizierung spezifischer, einzelner CpG-Loci ist diese Methode jedoch nur begrenzt geeignet. Aufgrund technischer Limitierungen ist bei dieser Methode die Analysierbarkeit vieler CpG-Loci nicht gegeben. Daher wurde in der vorliegenden Arbeit die Verifizierung von 450K-Ergebnissen hauptsächlich mittels BSPS durchgeführt, die valide Ergebnisse erbrachte.

4.2 Biologische Aspekte

In diesem Abschnitt werden die in dieser Arbeit gewonnenen Erkenntnisse der Analysen auf genetischer, epigenetischer und transkriptioneller Ebene hinsichtlich der biologischen Bedeutung für die Entstehung von Herzfehlbildungen wie dem HLHS diskutiert.

4.2.1 Identifizierung von Keimbahnmutationen bei syndromalen

CHDs und nicht-syndromalen Fällen mit HLHS

Die Mehrzahl der CHDs ist durch ein sporadisches Auftreten mit unbekannter Ätiologie charakterisiert [77]. Nach epidemiologischen Studien kann jedoch bei ca. 10 - 20 % der CHD-Patienten ein syndromaler Phänotyp, d.h. eine Assoziation zu genetischen Syndromen bzw. extrakardialen Anomalien, diagnostiziert werden [47, 66, 77, 408, 409]. Darüber hinaus lassen sich auch bei nicht-syndromalen CHD-Fällen Keimbahnmutationen in ursächlichem Zusammenhang mit der Entstehung der Herzfehlbildung feststellen [47]. Auch in dieser Arbeit ließen sich kausale Genvarianten bei Patienten mit einem syndromalen CHD als auch bei nicht-syndromalen HLHS-Fällen identifizieren.

In der vorliegenden Arbeit konnten kausale homozygote und *compound* heterozygote Mutationen in **TGDS** als Beispiel eines **syndrom-assoziierten CHDs** bei drei Familien mit Catel-Manzke-Syndrom als Beitrag zur Studie von Ehmke et al. (2014) [104] mittels WES-Analyse identifiziert und verifiziert werden (siehe 3.1.1). Die nach autosomal-rezessivem Vererbungsmodell auftretenden Mutationen im *TGDS*-Gen bei Patienten mit Catel-Manzke-Syndrom, welche in 40 % der Fälle eine Herzfehlbildung zeigen [103], umfassten die seltene Mutationen c.298G>T (p.Ala100Ser) und die bisher unbekannt Mutation c.294T>G (p.Phe98Leu). Erstere seltene Genvariante konnte in homozygoter Ausprägung bei zwei Patienten mit Catel-Manzke-Syndrom (aus Familie 1 und 2) identifiziert werden. Eine Homozygotie dieser Mutation, welche als *probably damaging* eingestuft ist, wurde bisher noch nicht beschrieben (SNP-Datenbankrecherche). Bei einer dritten Familie konnte darüber hinaus eine *compound* heterozygote Mutation c.298G>T/c.294T>G in *TGDS* beim Indexpatienten gezeigt werden. Da diese Mutationen seltene bzw. unbekannt Varianten mit wahrscheinlich protein-beschädigenden Einfluss darstellen, und darüber hinaus, zusammen mit vier weiteren *TGDS*-Mutationen, bei insgesamt sieben unabhängigen familiären Fällen mit Catel-Manzke-Syndrom identifiziert werden konnten [104], wurden diese Mutationen als kausale Genvarianten

in der Entstehung des Catel-Manzke-Syndroms angesehen. Das TGDS (TDP-Glucose 4,6-Dehydratase) ist ein evolutionär konserviertes Protein der *short-chain* Dehydrogenase/Reduktase (SDR)-Familie [410]. Beide Mutationen, p.Ala100Ser und p.Phe98Leu, liegen innerhalb der NAD-Bindedomäne des TGDS-Proteins (nach NCBI Conserved Domains Database, siehe 2.1.3.2) [104], sodass die katalytische Funktion dieser Domäne [410] durch diese Mutationen möglicherweise beeinträchtigt wird. Die spezifische Funktion des TGDS-Proteins beim Menschen ist noch unbekannt, jedoch hat das paraloge⁴⁴ Gen *UXS1*, eine Funktion in der Synthese von Glukosaminoglykanen [411], die Bestandteile von Proteoglykanen der extrazellulären Matrix (EZM) sind [412]. *TGDS*, welches zu 27 % dem *UXS1* auf Proteinlevel gleicht (*UXS1* Isoform 1, NP_001240804) [104], könnte somit in der Proteoglykansynthese involviert sein und Mutationen in diesem Gen könnten putativ die EZM-Ausbildung beeinträchtigen. Die EZM ist eine wichtige Komponente des Myokardiums und in allen Aspekten der Herzentwicklung involviert [413]. Störungen der Genexpression von EZM-Komponenten, oder auch Veränderungen von EZM-regulativen Signalwegen resultieren in kardialen Defekten, wie u.a. ASD und VSD, im Mausmodell [414 - 416]. Die in der vorliegenden Arbeit analysierten Familien mit Catel-Manzke-Syndrom umfassten auch einen Fall mit einem VSD. Der Entstehung dieses kardialen Defekts könnte somit möglicherweise eine Störung der EZM zugrundeliegen, begünstigt durch Mutationen im *TGDS*-Gen.

In ***NR2F2*** konnte in der vorliegenden Arbeit im Rahmen der Studie von Al Turki et al. (2014) [303] eine kausale *de novo* Splice-Donor-Mutation bei einem **nicht-syndromalen Fall mit HLHS** bei Analyse des Eltern-Index-Trios nachgewiesen werden (siehe 3.1.2.1). In der Studie von Al Turki et al. (2014) konnte gezeigt werden, dass Genvarianten in *NR2F2* in einem ursächlichen Zusammenhang mit AVSDs und LVOTO-Phänotypen (hier HLHS und CoA) stehen. Die *de novo* Splice-Donor-Mutation c.970+1G>A stellt eine bisher unbekannte Genvariante dar und führt wahrscheinlich durch Auslassen des dritten Exons zur Bildung eines trunkierten Proteins [303]. Einzelbasenpaar-Veränderungen in dem konservierten Dinukleotid GT der 5'-Splice-Donor-Site können die Bindung der U1 snRNA des Spliceosoms hemmen, sodass keine akkurate Exon-Inklusion stattfinden kann [417]. Demnach wäre in *NR2F2* (liganden-induzierbarer COUP-II TF), ein Teil der C-terminalen Liganden-Bindedomäne, welche die transkriptionelle Aktivität reguliert, von einer Trunkierung betroffen (nach ELM-Datenbank). Dies würde einen Verlust der TF-Funktion des Proteins bedeuten. Während der Organogenese zeigt der COUP-II TF eine hohe Expression in mesenchymalen

⁴⁴ Gene, die in der Evolution einer Spezies durch Duplikation eines Gens oder Chromosomenabschnitts, sowie anschließender Divergenz entstanden sind (Fitch, 1970; Review: Jensen, 2001)

Zellen [418, 419], sodass dieser TF möglicherweise eine wichtige Funktion in der endothelial-mesenchymalen Signalgebung hat [420]. Die endotheliale-mesenchymale Transition (EndMT) umfasst einen Prozess, in dem endotheliale Zellen ihre Zelltyp-spezifische Eigenschaften verlieren und in Zellen eines mesenchymalen Phänotyps umgewandelt werden [421]. Die EndMT ist für die Ausbildung von Herzklappen und der kardialen Septierung von wesentlicher Bedeutung [421 - 424]. Knockout-Experimente von *NR2F2* im Mausmodell ergaben atrioventrikuläre Septum- und Herzklappendefekte [420, 425], sodass *NR2F2* wahrscheinlich eine wichtige Funktion in der Herzentwicklung hat. In der Studie von Al Turki et al. (2014) [303] konnte zudem in einem Luziferase-Assay bei Analyse von sechs CDS-Varianten, die in *NR2F2* identifiziert wurden, eine signifikante Veränderung der TF-Aktivität von *NR2F2* an Ziel-Genpromotoren detektiert werden. Eine Haploinsuffizienz von *NR2F2* durch die detektierte *de novo* Splice-Donor-Mutation könnte somit zu einer aberranten Genexpression im Prozess der EndMT und somit zu einer fehlerhaften Herzklappenentwicklung [424] führen, sodass ein LVOTO-Phänotyp, wie das HLHS, entsteht.

4.2.1.1 Familien-Analyse von HLHS-Fällen: oligogene Vererbung mit reduzierter Penetranz und variabler Expressivität

Ein multifaktorielles, polygenes Vererbungsmodell, das das Auftreten multipler genetischer Varianten mit geringem Effekt und eine Interaktion von Umweltfaktoren umfasst, wurde lange Zeit als wesentliches Vererbungsmodell für CHDs angesehen [69, 75, 76]. Genomweite Analysen zu Multiplex-Familien mit CHDs konnten ein höheres Wiederholungsrisiko, als nach polygenem Vererbungsmodell zu erwarten wäre, zeigen [62, 66 - 68]. Wahrscheinlich liegen ein oligogenes Vererbungsmodell mit oder ohne reduzierter Penetranz, sowie das Auftreten von *de novo* Mutationen in der Entstehung von CHDs zugrunde [78, 236, 246, 304]. Auch die Beobachtung, dass identische Mutationen in CHD-assoziierten Genen ein variables Set verschiedenster Herzfehlbildungen auslösen [105, 426, 427], lässt vermuten, dass ein oligogener Vererbungsmechanismus, sowie ein möglicher Einfluss von Umweltfaktoren in der Entstehung von CHDs zugrunde liegt. Um mögliche Vererbungsmechanismen des HLHS zu untersuchen, wurden daher in dieser Arbeit genetische Analysen im familiären Kontext durchgeführt, deren Ergebnisse im Folgenden diskutiert werden.

In einer Familie mit zwei Geschwistern mit HLHS konnten die Gene *C15orf62*, *TNKS1BP1* und *PRF1* im Rahmen einer **WES-Studie zu Familien mit LVOTOs** (Hitz und Hoff et al., Publikation

in Vorbereitung (siehe 11)) als potenziell krankheitsrelevante Kandidatengene nach autosomal dominantem Vererbungsmodell mit reduzierter Penetranz identifiziert werden (siehe 3.1.2.2). Die Mutationen traten hierbei in den beiden betroffenen Geschwisterkindern mit HLHS und der phänotypisch gesunden Mutter auf, nicht jedoch bei dem gesunden Geschwisterkind und dem Vater. Der Erbgang der reduzierten Penetranz wurde als wahrscheinlichstes Vererbungsmodell basierend auf Generationen mütterlicherseits vermutet, da bei diesen Familienmitgliedern Totgeburten auftraten, welchen der Literatur nach als eine der häufigsten Todesursache CHDs [50], insbesondere zyanotische CHDs [306] wie das HLHS zugrunde liegen. Eine bisher unbekannte Stop-Mutation (c.757G>T (p.Glu253*)) konnte in *PRF1* detektiert werden, dessen Genprodukt Perforin 1 eine funktionelle Bedeutung im sekretorischen granulozyten-abhängigen Zelltod hat [428, 429]. Nach ELM-Datenbank wäre die Hälfte des Proteins von der Trunkierung betroffen, wobei wahrscheinlich die MACPF-Domäne (*membrane attack complex/perforin*), die der lytischen Membran-Porenbildung dient, funktionell deaktiviert würde. *In vitro*-Experimente an ventrikulären Kardiomyozyten zeigten, dass das Perforin eine Reduktion und Zerstörung dieser Zellen bewirkte [430]. Der programmierte Zelltod hat eine wesentliche Bedeutung in der Regulation der Zellproliferation und Eliminierung entarteter Zellen und konnte in verschiedenen Studien mit kardialen Anomalien in Verbindung gebracht werden [431 – 435]. Zwar ist das Perforin in infektionsbedingten Apoptose-Prozessen des Immunsystems involviert, eine Dysregulation dieser granulozyten-abhängigen Apoptose könnte jedoch möglicherweise auch in der Entstehung von Herzerkrankungen eine Rolle spielen.

Zusätzlich konnte in *C15orf62* eine bisher unbekannte Missense-Mutation, als *possibly damaging* eingestuft, detektiert werden. *C15orf62* stellt zwar ein Gen mit bisher unbekannter Funktion dar, jedoch konnte zuvor schon ein CHD-Fall mit wahrscheinlich pathogenen Ereignissen in der chromosomalen Region überlappend zu der von *C15orf62* in 15q15.1 gezeigt werden. Dieser CHD-Fall wies eine maternale Deletion in 15q15-q22.1 mit Diagnose einer Kraniosynostose, Missbildungen der Extremitäten, sowie einer TOF auf [436].

Als drittes potenziell krankheitsrelevantes Gen ließ sich in *TNKS1BP1* eine seltene Missense-Mutation nach dem Vererbungsmodell der reduzierten Penetranz feststellen. *TNKS1BP1*, das für ein Tankyrase 1-bindendes Protein kodiert, hat wahrscheinlich eine Funktion im deadenylierungs-abhängigen mRNA-*Decay* (Abbau der mRNA) nach Reactome-Datenbank [273] (siehe 2.1.3.2). In der Studie von Chang et al. (2005) [437] konnte mittels Knockdown im *Xenopus*-Modell gezeigt werden, dass Tankyrase 1, ein Enzym des DNA-Reperaturmechanismus, in der Assemblierung des Spindelapparats eine Rolle spielt. Dessen Knockdown führt zu einer Arretierung der Mitose [437, 438]. Das *TNKS1BP1*-Protein könnte

somit in Mitoseprozesse involviert sein. Ein Funktionsverlust durch die als *probably damaging* eingestufte Missense-Mutation in *TNKS1BP1* könnte möglicherweise Zellproliferations-Prozesse beeinflussen und in der Pathogenese des HLHS eine Rolle spielen.

Eine weitere Analyse im familiären Kontext umfasste in dieser Arbeit eine **SNP-Array-Analyse** einer Familie (Familie 6) mit einem nicht-syndromalen HLHS-Fall, um krankheitsassoziierte CNVs zu identifizieren (siehe 3.1.2.3). Ungefähr 3 – 10 % nicht-syndromaler CHDs liegen als ursächliche pathogene Variationen CNVs zugrunde [87, 91, 93]. Vererbte sowie *de novo* CNVs konnten zuvor in Assoziation zu LS-CHDs, inklusive dem HLHS, identifiziert werden [89]. Jedoch ließen sich in der vorliegenden Arbeit weder *de novo* CNVs, noch CNVs in Assoziation zu CHD-assoziierten Genen bzw. chromosomalen Regionen bei dem analysierten HLHS-Fall detektieren. Lediglich eine seltene paternal vererbte CNV konnte in 10q21.1 (0,29 Mb-Verlust) festgestellt werden. Zwar betrifft die CNV keine kodierende Region, jedoch konnte zuvor in *Linkage*-Analysen die chromosomale Region 10q21 - 10q23, welche mit der in dieser Arbeit detektierten CNV in 10q21.1 überlappt, in Assoziation mit der familiären DCM festgestellt werden [439]. Einige Familienmitglieder aus der Studie von Bowles et al. (1996) [439] wiesen zudem weitere kardiale Anomalien, wie MVP und Mitralklappeninsuffizienz, auf. CNVs in nicht-kodierenden Regionen können einen pathogenen Einflussfaktor in der Entstehung von CHDs darstellen, wenn diese CNVs potenziell *cis*-regulatorische DNA-Regionen betreffen [440] oder z.B. pathogene Positionseffekte bewirken [441]. Zwar konnte die CNV in 10q21.1 auch im phänotypisch gesunden Vater detektiert werden, es wird jedoch vermutet, dass viele CNVs eine inkomplette Penetranz und/oder variable Expressivität aufweisen [442]. Eine Kosegregation muss somit nicht zwingend eine Apathogenität bedeuten.

Ein oligogenes autosomal dominantes Vererbungsmodell mit reduzierter Penetranz und variabler Expressivität könnte somit einen möglichen Pathomechanismus für einige HLHS-Fälle darstellen. Eine reduzierte Penetranz ließ sich hierbei als wahrscheinlichstes Vererbungsmodell in der Familie 5 mit zwei HLHS-Fällen identifizieren. Auch eine variable Expressivität, d.h. ein unterschiedlicher Ausprägungsgrad der Symptome, könnte in einigen familiären Fällen mit LVOTOs zugrunde liegen. Familienmitglieder, wie hier die gesunden Eltern der mittels SNP-Array analysierten Familie 6, könnten asymptotische und damit undiagnostizierte, milde Ausprägungen von AoV-Defekten vorweisen [233, 237, 246]. Hierbei kann die Inzidenz von BAV (AoV-Defekt) bei Verwandten ersten Grades von Fällen mit LVOTOs wie dem HLHS bis zu 10 % betragen [233, 236]. Leider konnte jedoch keine Echokardiographie bei den jeweiligen Elternteilen durchgeführt werden, um diese Hypothese zu bestätigen. Zusammenfassend ließen sich die Beobachtungen und Hypothesen dieser Arbeit zum Vererbungsmodell des HLHS mit

reduzierter Penetranz und/oder variabler Expressivität auch zuvor in anderen Studien zu familiären BAV-Fällen und weiteren LVOTO-Phänotypen zeigen [78, 236, 246, 304]. Ein solcher Vererbungsmechanismus könnte somit ein mögliches Modell für einen Anteil von HLHS-Fällen darstellen.

4.2.2 Potenziell krankheitsrelevante Genvarianten im Herzgewebe von Patienten mit HLHS

Im IAS-Gewebe von neun Patienten mit HLHS konnten in dieser Arbeit insgesamt 685 Gene mit seltenen bzw. bisher unbekanntem funktionellen Genvarianten detektiert werden, von denen 56 Varianten LOF-Mutationen darstellten (siehe 3.2.1.1). Die Herzgewebeproben der HLHS-Patienten wiesen in mehr als einem Drittel dieser 685 Gene funktionelle Varianten in Genen auf, die nach ExAC-Datenbank in > 60.000 Individuen eine besonders geringe Häufigkeit funktioneller Varianten und somit eine geringe Mutationstoleranz nach RVIS-Scoring-System [264] zeigen, sodass diese Genvarianten potenziell einen pathogenen Einfluss in der Entstehung dieser Herzfehlbildung haben könnten. Innerhalb der 685 Gene mit Mutationen bei HLHS-Fällen zeigten Gene mit bekannter Assoziation zu Herzfehlbildungen (Herz-Genliste) mit einem RPS-Median < 20 eine besonders geringe Mutationstoleranz (siehe 3.2.1.1.2 Abbildung 22), sodass zwar lediglich 27/685 Gene mit einem bekannten „Herzphänotyp“ detektiert wurden, diese jedoch mit erhöhter Wahrscheinlichkeit eine Pathogenität der Genvarianten beinhalteten. Eine Gennetzwerkanalyse mittels STRING-Software ergab innerhalb der 685 Gene Interaktionsnetzwerke, die in Entwicklungsprozessen, wie Notch- und Wnt-Signalwegen, sowie in der (post-)transkriptionellen Regulation involviert sind (siehe 3.2.1.1.2 Abbildung 23). Mutationen in Genen der transkriptionellen Regulation, wie bspw. TFs der Herzentwicklung, und der Signaltransduktion von Entwicklungs-*Pathways* sind auch in der Literatur als maßgeblich betroffene Gen-Kategorien im Zusammenhang mit nicht-syndromalen CHDs beschrieben [47]. Betrachtet man aus der Literatur bekannte, mit Herzfehlbildungen assoziierte Gene, so waren 17/685 Gene mit einem kardiovaskulären Phänotyp im Mausmodell assoziiert und 10/685 Gene konnten zuvor bei Fällen mit kardiovaskulären Erkrankungen identifiziert werden (siehe 3.2.1.1.2 Tabelle 24).

Überlappende Gene mit einem kardialen Phänotyp im Mausmodell

Innerhalb der Gene mit einem „Maus-Phänotyp“ konnte eine Missense-Mutation in *HAND1* bei einem HLHS-Patienten identifiziert werden. *HAND1* stellt einen wesentlichen TF in der frühen Herzentwicklung dar und ist in der Regulation der kardialen Ventrikel-Expansion involviert [443]. Knockout-Mäuse wiesen Störungen des Looping-Prozesses und eine hochgradige Hypoplasie des LVs und Ausflusstrakt-Anomalien [309, 310], sowie Defekte des LVs und eine dysregulierte ventrikuläre Genexpression auf [443]. Ähnliche Prozesse sind auch bei der Entwicklung eines HLHS betroffen, sodass Genvarianten in *HAND1* in der Pathogenese des HLHS eine Rolle spielen könnten. Weitere in dieser Arbeit bei HLHS-Patienten identifizierte Gene mit einem kardiovaskulären Phänotyp im Mausmodell umfassten *IFT88* und *PTCH1*, welche beide in Hedgehog-Signalwegen eine Rolle spielen. Diese Signalwege haben in der endokardialen Spezifikation und Differenzierung eine wesentliche Funktion [444 - 447]. In *IFT88* konnte eine Frameshift-Mutation mit vorzeitigem Stopkodon (c.1508delT (p.F504Lfs*17)) nachgewiesen werden, welche eine Trunkierung der Hälfte des Proteins nach ELM-Datenbank und somit wahrscheinlich einen Funktionsverlust bewirkt. Ein Knockout von *IFT88* im Mausmodell zeigte, dass *IFT88* in der frühen Kardiogenese eine wesentliche Funktion in der Koordination des Hedgehog-Signalweges hat [308]. Das Gen *PTCH1*, in dem eine Missense-Mutation nachgewiesen werden konnte, kodiert für einen Rezeptor von Hedgehog-Proteinen [448] und ist somit direkt in der Hedgehog-Signalgebung involviert. Der Funktionsverlust im Mausmodell von *SHH* (Sonic Hedgehog) resultiert in ventrikulärer Hypoplasie, Septationsdefekten und Ausflusstraktverkürzung [449, 450]. Mutationen in Genen des Hedgehog-Signalweges, wie in dieser Arbeit detektiert wurde, könnten einen möglichen pathogenen Mechanismus in der Entstehung des HLHS darstellen.

Überlappende Gene mit bekannten kardiovaskulären Erkrankungen

Mittels der WES-Analyse nachgewiesene Mutationen in Genen, die zuvor bei Fällen mit kardiovaskulären Erkrankungen identifiziert werden konnten, umfassten Gene, die bei syndrom-assoziierten sowie nicht-syndromalen CHDs und auch bei Kardiomyopathie-Fällen identifiziert wurden (siehe 3.2.1.1.2 Tabelle 24). Da es sich bei den in dieser Arbeit analysierten HLHS-Patienten um isolierte, nicht-syndromale Fälle handelte, wird im Folgenden auf bekannte Gene nicht-syndromaler Fälle eingegangen.

In *NAA15*, dessen Genprodukt (N-Acetyltransferase) in Angiogeneseprozessen involviert ist [262], konnte jeweils eine seltene und eine bisher unbekannte heterozygote Missense-Mutation in HLHS_12 und HLHS_21/22 nachgewiesen werden. Die unbekannte Genvariante c.945G>C (p.Arg315Ser), welche bei HLHS_21/22 detektiert werden konnte, wird nach Polyphen-2 zudem als *probably damaging* eingestuft. In der Studie von Zaidi et al. (2013) [74] ließen sich bei zwei Patienten mit CHDs *de novo* LOF-Mutationen in *NAA15* nachweisen, sodass eine Haploinsuffizienz dieses Gens potenziell in einem ursächlichen Zusammenhang mit der Entwicklung eines HLHS steht.

Des Weiteren konnte bei Patient HLHS_15 in *NOTCH1*, das für einen Transmembranrezeptor kodiert, eine bisher unbekannte Frameshift-Mutation mit vorzeitigem Stopkodon detektiert und verifiziert werden. Der Notch-Signalweg stellt einen evolutionär hochkonservierten Signalweg dar, welcher in der Regulation von Zellschicksals- und Differenzierungsprozessen der Embryogenese, inklusive der kardialen Entwicklung, eine wesentliche Bedeutung hat [451, 452]. Die Aktivierung des Notch1-Rezeptors bewirkt eine Spaltung und nukleäre Translokation der intrazellulären Domäne, welche die Transkription von Zielgenen der Kardiogenese steuert [451]. Verschiedene Tiermodelle zeigten, dass *NOTCH1* in der vaskulären Morphogenese involviert ist [453], sowie in der frühen Valvulogenese eine Rolle spielt [223]. Die Notch1-Signalgebung ist von wesentlicher Bedeutung in der EndMT [223, 224, 454 – 457], einem wichtigen Prozess der Herzklappenentwicklung [421 - 424]. Heterozygote *NOTCH1*-Mutationen ließen sich bisher bei einem weiten Spektrum von LS-CHDs, wie BAV [225, 458 - 461], AoV-Stenose und CoA [225] sowie dem HLHS identifizieren [196, 225, 226]. Somit ist eine Anreicherung von Missense- als auch LOF-Mutationen in *NOTCH1* bei der Entstehung von LS-CHDs wie dem HLHS, zu postulieren (siehe Abbildung 59). Kürzlich konnte zudem in einer WES-Studie von 1823 CHD-Familien (Sifrim und Hitz et al., 2016⁴⁵), zu der im Rahmen dieser Arbeit beigetragen wurde, eine Anreicherung von vererbten LOF-Mutationen und *de novo* Mutationen in *NOTCH1*, sowie eine signifikante Anreicherung von Genen des *NOTCH1-Pathways* bei Patienten mit CHDs detektiert werden. Die in dieser Arbeit identifizierte LOF-Mutation bei HLHS_15 führt auf Proteinlevel zu einer Trunkierung des gesamten Proteins, da die komplexe INDEL-Mutation c.408_414delATCGTGCinsGGGGCTGAC (p.Ser137Glyfs*140) direkt am N-Terminus des 2555 aa großen Proteins (NP_060087.3) in Exon 4/34 auftritt (Abbildung 59), was einen kompletten Funktionsverlust des Proteins bedeuten würde. Die von der Frameshift-Mutation

⁴⁵ Dokument derzeit in Revision, eingereicht bei Nature Genetics (siehe Abschnitt 11)

betroffenen Aminosäuren sind zudem in den verschiedenen Spezies hochkonserviert (Abbildung 59 „Multiz Alignments of 100 Vertebrates“).

Die Position der Frameshift-Mutation von HLHS_15 liegt sehr nahe zu den beiden *de novo* LOF-Mutationen in Exon 4, die in der Studie von Iacone et al. (2012) [196] bei zwei isoliert auftretenden HLHS-Fällen identifiziert werden konnten (siehe Übersicht *NOTCH1*-Mutationen Abbildung 59). Nach Mutation Taster [268] wurden letztgenannte *de novo* LOF-Mutationen als *disease causing* eingestuft [196]. Die in der Studie von Iacone et al. (2012) [196] identifizierte Frameshift-Mutation p.Lys136Asnfs*141 endet hierbei auf demselben vorzeitigen Stopkodon, wie die in der vorliegenden Arbeit identifizierte Frameshift-Mutation bei HLHS_15 (siehe Abbildung 59). Somit konnte funktionell dieselbe Mutation zuvor auch schon bei einem weiteren HLHS-Fall identifiziert werden. Eine Untersuchung von Familienmitgliedern des inzwischen verstorbenen Patienten HLHS_15 war leider nicht möglich, sodass diese Mutation nicht in familiären Kontext analysiert werden konnte, um eine mögliche *de novo* oder Herzwegewebe-spezifische Ausprägung der *NOTCH1*-Mutation zu bestätigen. *NOTCH1* hat einen RPS-Wert von 0,33, sodass die Mutationstoleranz gegenüber funktionellen Varianten geringer als die von 99 % anderer Gene des Genoms ist [264]. Nach ExAC-Datenbank ist zudem in dem *NOTCH1*-Gen bisher lediglich eine seltene LOF-Mutation (Splice-Donor-Mutation, detektiert in 1/63728 Individuen) bekannt. Somit handelt es sich bei der LOF-Mutation bei HLHS_15, auch in Anbetracht der Tatsache, dass funktionell dieselbe Mutation zuvor schon bei einem anderen HLHS-Fall als pathogene Variante identifiziert wurde [196], um eine wahrscheinlich kausale Mutation.

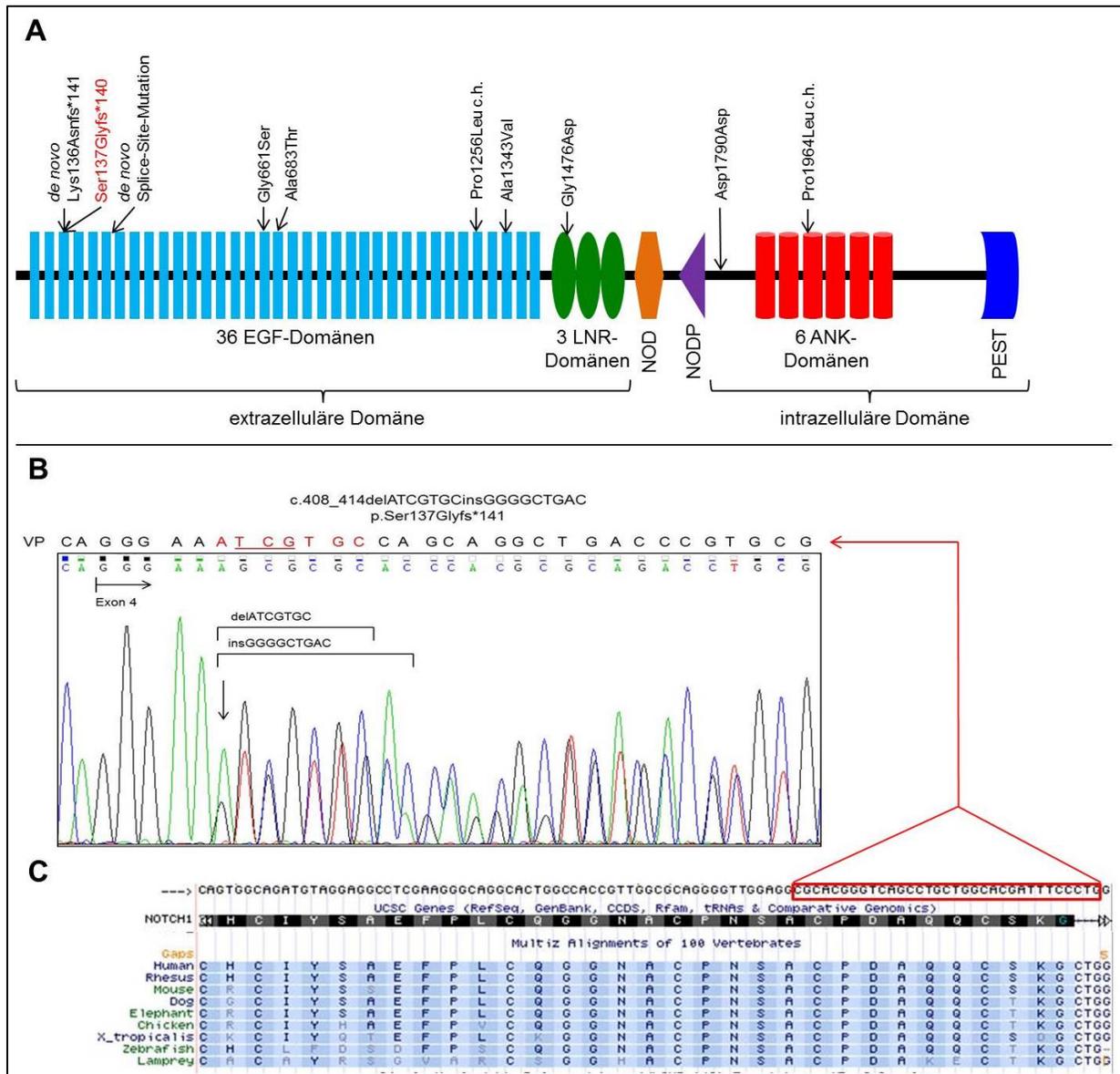


Abbildung 59 Übersicht von *NOTCH1*-Mutationen, die bei Patienten mit HLHS identifiziert wurden (A) und Elektropherogramm der *NOTCH1*-Mutation bei HLHS_15 (B) mit Darstellung im UCSC-Genome Browser (C). A: Schematische Darstellung des *NOTCH1*-Proteins (NM_017617.3, ENSP0000277541.6) mit Angaben zu bekannten Domänen, erstellt nach ELM-Datenbank. Bisher bekannte Genvarianten in *NOTCH1*, die als wahrscheinlich pathogene Varianten nach Studien von McBride et al. (2008) [225], lascone et al. (2012) [196], und Theis et al. (2015) [226] bei Patienten mit HLHS identifiziert wurden, sind in der schematischen Proteindarstellung markiert und als Mutation auf Proteinebene beschriftet. Bis auf zwei *de novo* Mutationen und eine *compound* heterozygote Mutation (c.h.) sind alle weiteren Mutationen heterozygot ausgeprägt. Die in der vorliegenden Arbeit beim Patienten HLHS_15 identifizierte Mutation p.Ser137Glyfs*140 ist in roter Schrift dargestellt. EGF: *epidermal growth factor-like*, LNR: Notch/Ln12-*Repeats*, NOD/NODP: Regionen in Notch-Proteinen mit unbekannter Funktion, ANK: Ankyrin-*Repeats*, PEST: Prolin-, Glutamat-, Serin-, Threonin-reich B: Elektropherogramm der Sanger-Sequenzierung der *NOTCH1*-Mutation c.408_414delATCGTGCinsGGGGCTGAC (p.Ser137Glyfs*140) im IAS-Gewebe von HLHS_15. Der Beginn der Frameshift-Mutation in Exon 4 ist durch einen Pfeil markiert. Oberhalb des Elektropherogramms ist jeweils die Sequenz des Vorwärts-Primers (VP) dargestellt. Rote Buchstaben stellen die mutierten Base dar, unterstrichene Buchstaben die erste mutierte Kodonsequenz der Aminosäuren. C: Darstellung des Elektropherogramm-Ausschnitts im UCSC-Genome Browser (rot markiert), als komplementär-reverse Sequenz des *NOTCH1*-Gens, welches auf dem Minusstrang der DNA liegt. Die 3' nachfolgenden Aminosäuren der Position der Frameshift-Mutation sind auch in anderen Spezies hochkonserviert (siehe Multiz Alignments of 100 Vertebrates).

Das Auftreten von *NOTCH1*-Mutationen mit unterschiedlichem Schweregrad (BAV als milde Form bis zu HLHS als schwerwiegende CHD) [196, 225, 226, 458 - 461], sowie eine reduzierte Penetranz von *NOTCH1*-Mutationen mit hoher Variabilität intra- und interfamilärer phänotypischer Expression bei Fällen mit LS-CHDs [225, 458] lassen einen Einfluss zusätzlicher genetischer Komponenten als auch mögliche epigenetische Faktoren in der Entstehung des HLHS vermuten. Eine unterschiedliche Mutationslast von *NOTCH1* und synergistische Mutationen weiterer Komponenten des Notch-Signalweges könnten das Ausmaß der Beeinträchtigung der Notch1-Signalgebung beeinflussen und somit den unterschiedlichen Schweregrad von LS-CHDs erklären.

4.2.2.1 Somatische Mutationen beim HLHS

Um mögliche Herzgewebe-spezifische Mutationen zu detektieren, wurde in dieser Arbeit eine kombinierte Keimbahn- und somatische Analyse in einer Familie mit HLHS durchgeführt (siehe 3.2.1.2). Die Hypothese bestand darin, dass möglicherweise ursächliche somatische Mutationen in dem betroffenen Gewebe, also dem Herzgewebe, auftreten könnten, die jedoch nicht in der Keimbahn vorliegen und somit in einem Mosaizismus resultieren. Wie in Abschnitt 4.1.2.1 diskutiert, war aufgrund von Limitationen der Sanger-Sequenzierungstechnologie die Verifizierung von zuvor mittels WES-Technologie im IAS-Gewebe detektierten Missense-Mutationen in *ANKRD11* und *BHLHA15* bisher nicht möglich. Es sind daher *Deep-Sequencing* NGS-Analysen geplant, welche jedoch nicht mehr Bestandteil dieser Arbeit sind. Da die Mutationen jeweils mittig innerhalb der überlappenden *Reads* und zudem jeweils in FP- und RP-*Reads* detektierbar waren, wurden Sequenzier-Artefakte als unwahrscheinlich erachtet. Beide Mutationen wurden als *possibly damaging* durch Polyphen-2 eingestuft und stellten nach ExAC-Datenbank bisher unbekannte Varianten dar, sodass es sich möglicherweise um krankheitsrelevante Mutationen handelt.

ANKRD11 kodiert für das Ankyrin Repeat Domain 11-Protein, welches einen Inhibitor der Liganden-abhängigen Aktivierung der Transkription darstellt [462]. Heterozygote Mutationen in *ANKRD11* konnten in ursächlichem Zusammenhang mit dem KBG-Syndrom identifiziert werden, das durch geistige Behinderung, Minderwuchs, Skelettanomalien, sowie in einigen Fällen auch CHDs, wie BAV und partieller AVSD charakterisiert ist [463]. Innerhalb der syndrom-assoziierten CHD-Kohorte der WES-Studie von Sifrim und Hitz et al. (2016⁴⁶), an der sich im Rahmen dieser

⁴⁶ Dokument derzeit in Revision, eingereicht bei Nature Genetics (siehe Abschnitt 11)

Arbeit beteiligt wurde, konnte eine genomweit signifikante Anreicherung von *de novo* Mutationen in *ANKRD11* gefunden werden. Darüber hinaus stellt *ANKRD11* mit einem RPS-Wert von 0,62 ein Gen mit besonders geringer Mutationstoleranz nach RVIS-Scoring-System [264] dar, sodass Mutationen in diesem Gen einen möglichen pathogenen Faktor in der Entstehung von CHDs darstellen.

BHLHA15, das für ein Basic Helix-Loop-Helix-Protein kodiert, stellt einen transkriptionellen Repressor von *MYOD1* dar, einem TF mit funktioneller Bedeutung bei der Differenzierung myogener Stammzellen [464]. *MYOD1* zeigt nach Pathway Interaction Database [271] (siehe 2.1.3.2) eine Assoziation mit dem Notch-abhängigen HES/HEY- *Pathway*, sodass *BHLHA15* indirekt in der Notch-Signalgebung involviert sein könnte. Mutationen in letztgenanntem Signalweg lassen, wie in Abschnitt 4.2.2 diskutiert wurde, einen möglichen Pathomechanismus in der Entstehung des HLHS vermuten. Somatische Mutationen in kardialen Vorläuferzellen der frühen Herzentwicklung in *BHLHA15* könnten hierbei Veränderungen des Notch-Signalweges hervorrufen, sodass eine gestörte EndMT [223, 224, 454 – 457] und somit eine abnorme Herzklappenentwicklung [424], wie dies beim HLHS der Fall ist, resultiert.

Solche somatischen Mutationen könnten in unterschiedlichen Phasen der Herzentwicklung in Subpopulationen kardialer Zellen oder bestimmten Herzgeweberegionen als Mosaik auftreten. Eine bei konstitutioneller Ausprägung theoretisch letale Auswirkung auf den sich entwickelnden Fetus würde durch den somatischen Mosaizismus maskiert werden [465]. Je nach betroffener Herzgeweberegion und Zeitpunkt des Auftretens solcher Aberrationen in der Herzentwicklung könnte ein unterschiedlicher Schweregrad des kardialen Phänotyps resultieren. In Anbetracht dessen, dass ein gehäuftes Auftreten von BAV, als eine milde LVOTO-Form, mit bis zu 10 % bei Familienmitgliedern von HLHS-Fällen festzustellen ist [236 – 238], könnten in der Entstehung eines HLHS somatische Ereignisse mit konsekutiven kardialen Defekten unterschiedlichem Schweregrads zugrunde liegen. Solche Mosaik-Mutationen im Herzgewebe könnten einen möglichen Mechanismus für das Auftreten von inkompletter Penetranz beim HLHS (siehe 4.2.1.1) darstellen, deren Nachweis jedoch *Deep-Sequencing*-Technologien erfordert (siehe 4.1.2.1).

4.2.3 Epigenetische Aspekte in der Analyse von Herzgewebe

Um mögliche pathogene Faktoren im Herzgewebe auch auf epigenetischer Ebene zu untersuchen, wurde in dieser Arbeit DNA-Methylierungsanalysen mittels 450K-Array an Herzgewebe von Patienten mit CHDs, insbesondere mit HLHS, durchgeführt. Die Ergebnisse der verschiedenen Teilprojekte werden im Folgenden diskutiert.

4.2.3.1 Gewebespezifische DNA-Methylierungsprofile verschiedener

Herzregionen

In dieser Arbeit wurde die DNA-Methylierung in Herzgewebeproben verschiedener Regionen des gesamten Herzens von Patienten mit CHDs untersucht, um zu analysieren, ob die unterschiedlichen kardialen Regionen differenzielle DNA-Methylierungsmuster aufwiesen. Diese Herz-Mapping-Analyse ergab signifikant differenzielle DNA-Methylierungsprofile von Gewebeproben des Atriums, ventrikulären Proben sowie Gewebeproben von Gefäß(ähnlichen)-Strukturen (siehe 3.2.2.2). Diese drei Kategorien zeigten bereits in der unsupervidierten Analyse eine klare Segregation in unterschiedliche Gruppierungen (siehe 3.2.2.2 Abbildung 32). Daher wurden diese speziellen Gewebe-Signaturen in weiteren Teilprojekten analysiert (siehe 3.2.2.3 und 3.2.2.4), deren Ergebnisse in den nachfolgenden Abschnitten diskutiert werden.

Die unterschiedlichen Herzregionen zeigten insgesamt eine signifikante Anreicherung von RDMR- (OR = 4,65) und Enhancer-assoziierten (OR = 3,25) CpG-Loci innerhalb der DM-CpGs. RDMRs, d.h. reprogrammierungsspezifische DMRs, umfassen CpG-Loci der epigenetischen Remodellierung und haben eine zentrale Bedeutung in der somatischen Zellreprogrammierung [173]. Letzteres umfasst den Prozess der Reprogrammierung einer somatischen Zelle, um einen pluripotenten Status ähnlich einer embryonalen Stammzelle zu erhalten, d.h. die Herstellung induzierbarer pluripotenter Stammzellen (iPSCs) [466]. Nach Studien von Doi et al. (2009) könnten RDMRs von CpG-Insel-Shores (um 2 kb angrenzende DNA-Region von CpG-Inseln (CGIs)) kritische epigenetische Ziel-Loci der Determinierung des Zellschicksals darstellen. Es lagen 64 % (98/154) der RDMR-assoziierten DM-CpG-Loci innerhalb von CGI-Shores, sodass mehr als die Hälfte der identifizierten RDMR-assoziierten DM-CpG-Loci möglicherweise in Methylierungsprozessen der Zellschicksals-Definition unterschiedlicher Herzregionen involviert ist. Die differenzielle DNA-Methylierung solcher CGI-Shores zeigt eine hohe Korrelation mit einer gewebespezifischen Genexpression [467],

sodass die detektierte herzregionspezifische differenzielle DNA-Methylierung potenziell in der Regulation differenzieller Genexpressionsmuster verschiedener Regionen des Herzens involviert ist. Die signifikante Anreicherung Enhancer-assoziiierter DM-CpGs lässt vermuten, dass insbesondere genregulatorische Elemente in der herzregionspezifischen differenziellen DNA-Methylierung eine Rolle spielen. In der GO-Analyse fand sich zudem eine signifikante Anreicherung von Genen, die in der Embryonalentwicklung eine Rolle spielen. Zusammenfassend lässt sich daher annehmen, dass unterschiedliche Herzregionen besonders in Entwicklungsgenen durch ein differenzielles DNA-Methylierungsprofil, speziell von genregulatorischen und Zell-Determinations-assoziierten CpG-Loci, charakterisiert sind.

4.2.3.1.1 Differenzielles DNA-Methylierungsprofil und potenzielle epigenetische

Marker atrialer und ventrikulärer Herzgewebeproben

In dieser Arbeit konnte ein signifikant differenzielles DNA-Methylierungsprofil von atrialem und ventrikulärem Herzgewebe mit Methylierungsdifferenzen von $\Delta\beta > 30\%$ nachgewiesen werden (siehe 3.2.2.3). Zwei Drittel der identifizierten DM-CpGs zeigten eine Hypermethylierung des Ventrikelmyokards im Vergleich zu atrialem Gewebe, wobei eine signifikante Anreicherung von RDMR- (OR = 4,24) und Enhancer-assoziierten (OR = 3,63) DM-CpG-Loci gezeigt werden konnte. Die kardialen Gene *IRX4* (Iroquois Homeobox 4) und *FGFR2* (Fibroblast Growth Factor Receptor 2) zeigten hierbei multiple DM-CpG-Loci pro Gen (3 CpG-Loci) mit einer signifikanten Hypermethylierung in Ventrikelgewebe verglichen mit Atriumgewebe (siehe 3.2.2.3.2 Abbildung 38) und könnten daher epigenetische Biomarker dieser gewebespezifischen DNA-Methylierung darstellen.

Homeodomäne-TFs der Iroquois-Familie, wie *IRX4*, sind von wesentlicher Bedeutung in der Regulation von Entwicklungsprozessen [468] und *IRX4* stellt einen frühen Marker ventrikulärer Vorläuferzellen mit einer vermehrten Expression in Ventrikelgewebe während sämtlicher kardialer Entwicklungsstadien in Maus, Huhn und *Xenopus* dar [469 - 471]. Eine *IRX4*-Defizienz im Mausmodell führt zu einer abnormalen ventrikulären Genexpression und Kardiomyopathie [472], sodass dieses Gen eine kritische Rolle in der ventrikulären Differenzierung im Zuge der Kardiogenese spielt. Eine solche „Ventrikelspezifität“ ließ sich auch in der vorliegenden Arbeit identifizieren, indem in *IRX4* eine signifikante Hypermethylierung von intragenischen CGIs innerhalb des Genkörpers von ventrikulären Gewebeproben im Vergleich zu atrialen Gewebeproben nachgewiesen werden konnte. Gewebespezifische intragenische

CGI-Methylierungsmuster ließen sich zuvor schon in verschiedenen Studien detektieren [473, 474]. Nach Illingworth et al. (2010) [473] könnten diese intragenischen CGIs „*orphan promoters*“ repräsentieren, deren transkriptionelle Aktivität in frühen Entwicklungsstadien von Bedeutung ist. Eine differenzielle DNA-Methylierung dieser Loci in atrialem und ventrikulärem Gewebe könnte somit in direktem Zusammenhang mit einer herzregionspezifischen Genexpression von *IRX4* stehen.

Die FGF-Signalgebung spielt eine wichtige Rolle in der frühen Kardiogenese, was mithilfe verschiedener Modellorganismen nachgewiesen werden konnte [475 - 481]. Analysen im Mausmodell lassen vermuten, dass *FGFR2* in der Ausflusstraktseptierung und der ventrikulären Morphogenese eine wichtige Rolle spielt [482]. In *FGFR2* ließ sich in der vorliegenden Arbeit eine signifikante Hypermethylierung in ventrikulärem Gewebe im Vergleich zu atrialem Gewebe nachweisen, wobei zwei der drei DM-CpG-Loci in Assoziation mit CGI-Shores vorlagen. Da in der Studie von Irizarry et al. (2009) [467] eine hohe Korrelation von differenzieller DNA-Methylierung von CGI-Shores mit einer gewebespezifischen Genexpression gezeigt werden konnte, könnte neben *IRX4* somit auch *FGFR2* einen epigenetischen Marker der Gewebespezifität des Herzens darstellen.

Die identifizierten DM-CpG-Loci zwischen Atrium- und Ventrikelgewebe wiesen insgesamt eine Anreicherung von Entwicklungsgenen des kardiovaskulären Systems auf. Dies lässt zum einen die Feststellung zu, dass insbesondere Gene der Kardiogenese vornehmlich in genregulatorischen Elementen eine differenzielle DNA-Methylierung zwischen atrialem und ventrikulärem Herzgewebe zeigen. Zum anderen könnte die vermehrte Hypermethylierung der untersuchten ventrikulären Gewebeproben einen möglichen Einflussfaktor in der Entstehung der entsprechenden CHD-Typen darstellen. Die Ventrikelproben umfassten in dieser Analyse LV/RV-Proben je eines HLHS-TGA- und LV-An-Patienten und eine LV-Probe eines DCM-Patienten. Es zeigte hierbei hauptsächlich der LV pathologische Veränderungen, mit (1) einer Hypoplasie bei HLHS-TGA, (2) einer dünnen, dyskinetischen LV-Wand mit fibrösem Gewebe und nekrotischem Myokard beim LV-An und (3) einer Dilatation und Strukturveränderung des Myokards bei der DCM. Eine Hypermethylierung von CpG-Loci in Enhancer-Regionen, die eine signifikante Anreicherung innerhalb der DM-CpG-Loci zeigten (s.o.), könnte in einer aberranten Genexpression im LV resultieren. Die DNA-Methylierung von Enhancer-Regionen könnte die Bindung von Transkriptions-Aktivatorproteinen an den Enhancer beeinträchtigen, sodass die transkriptionelle Aktivierung von Genen gehemmt würde. In der Studie von Blattler et al. (2014) [176] konnte gezeigt werden, dass ein Verlust der DNA-Methylierung in einer Aktivierung von Enhancer-Regionen mit Veränderungen der Genexpression einhergeht.

Ebenso konnten Wiench et al. (2011) [483] eine positive Korrelation von Enhancer-Demethylierung und TF-Bindung feststellen. Eine Hypermethylierung von Enhancer-CpGs von Herzentwicklungsgenen in ventrikulärem Gewebe könnte somit eine dysregulierte Genexpression und konsekutiv pathologische LV-Entwicklung zur Folge haben.

Atriales und ventrikuläres Herzgewebe unterscheidet sich auf morphologischer und funktionaler Ebene, insbesondere hinsichtlich elektrophysiologischer und kontraktiver Eigenschaften [484]. Bisher wurden jedoch nur wenige Gene in Zusammenhang mit Unterschieden von atrialen und ventrikulären Kardiomyozyten identifiziert. Dies umfasst u.a. eine differenzielle Genexpression einiger Strukturgene [484] oder auch TFs, wie u.a. *IRX4* [469, 470, 485]. Letztgenanntes Gen konnte auch in dieser Arbeit auf epigenetischer Ebene mit einer gewebespezifischen differenziellen DNA-Methylierung zwischen atrialem und ventrikulärem Gewebe identifiziert werden, sodass eine differenzielle DNA-Methylierung in *IRX4* einen möglichen Mechanismus in der Regulation der gewebespezifischen Expression darstellen könnte. Eine differenzielle DNA-Methylierung als distinkte epigenetische Signatur ventrikulärer Kardiomyozyten konnte zudem zuvor in einer Studie von Xu et al. (2012) [486] in der Differenzierung pluripotenter Stammzellen gezeigt werden, sodass die DNA-Methylierung einen wichtigen Faktor in der Differenzierung verschiedener Herzregionen darstellt.

Zusammenfassend könnte somit eine differenzielle DNA-Methylierung von *IRX4* und *FGFR2* einen potenziellen epigenetischen Marker der Gewebespezifität von atrialem und ventrikulärem Herzgewebe darstellen.

4.2.3.1.2 Muskel- und Gefäß-Signatur und mögliche Kandidatengene in der Entstehung einer DCM

In dieser Arbeit konnte bei der Analyse von neun Herzgewebeproben eines DCM-Patienten (IAS, LA, RA, LV, MV, TV, PA, Aorta und AoV) ein signifikant differenzielles DNA-Methylierungsprofil von muskulärem Gewebe des Herzens im Vergleich zu Gefäßstrukturen mit > 30 %igen Methylierungsdifferenzen detektiert werden (siehe 3.2.2.4). Eine Analyse der globalen DNA-Methylierung verschiedener Herzregionen von DCM-Patienten ist in der Literatur bisher noch nicht bekannt und wurde somit im Rahmen der vorliegenden Arbeit erstmalig durchgeführt. Bereits in der Gesamt-CpG-Korrelationsmatrix (Abbildung 40) und in der unsupervidierten PCA (Abbildung 41) ließ sich eine deutliche Segregation von „Muskel“- und „Gefäß“-Proben feststellen. Neben den beiden Gefäß-Proben PA und Aorta segregierte auch die

AoV-Probe zusammen mit diesen Gefäß-Proben. Dies lässt sich dadurch erklären, dass bei der Gewebeexzision der Bereich des Aortenklappenrings, dem Beginn der Aorta *ascendens* [248, 320], entnommen wurde (siehe 2.1.1.1). Dieser enthält, genauso wie die Wandschichten *Tunica media* und *Tunica adventitia* der Gefäße, kollagenes Bindegewebe [248]. Die histologische Ähnlichkeit dieser Proben war somit durch ein gemeinsames Methylierungsprofil erkennbar. Die atrioventrikulären Herzklappen MV und TV zeigten eine Gruppierung mit muskulären Gewebeproben (LA, RA, LV). Bei diesen Herzklappen-Proben wurde die Gewebeexzision im Bereich der Klappenaufhängung durchgeführt (siehe 2.1.1.1). Die Klappensegel der atrioventrikulären Herzklappen sind durch *Chordae tendineae* an der Ventrikelwand über Papillarmuskel verbunden [249], sodass hier bei der Gewebeexzision auch muskuläre Anteile anfallen. Dies erklärt die Segregation der Herzklappen-Proben zusammen mit den muskulären Gewebeproben. Drei Gene, *GRIK2* (11 CpGs), *DIP2C* (10 CpGs) und *PRDM16* (9 CpGs) wiesen mit 9 – 11 CpGs eine besonders hohe Anzahl DM-CpGs pro Gen auf (siehe 3.2.2.4 Abbildung 43) und werden daher im Folgenden diskutiert.

In *GRIK2*, das für einen Glutamat-Rezeptor kodiert, lag eine signifikante Hypomethylierung im Herzmuskelgewebe im Vergleich zu Gefäßstrukturen vor, wobei vornehmlich CpG-Loci (9/11 CpGs) innerhalb von CpG-Inseln 200 bp *upstream* des Transkriptionsstarts (TSS) und innerhalb des ersten Exons betroffen waren. Verschiedene Studien konnten eine negative Korrelation der DNA-Methylierung von Promotorregionen, und damit DNA-Regionen *upstream* der TSS, mit der Expression von *downstream* gelegenen Genen zeigen [165, 166, 487 - 490]. Die nachgewiesene Hypomethylierung dieser CpG-Loci von *GRIK2* ließe somit eine transkriptionelle Aktivität dieses Gens in muskulärem Gewebe und eine verminderte Genaktivität in Gefäßstrukturen vermuten. Für eine eindeutige Aussage hinsichtlich des Zusammenhangs der gewebespezifischen *GRIK2*-Methylierung und -Expression sind zukünftig funktionelle Studien geplant.

DIP2C (Disco-Interacting Protein 2 Homolog C), ein proteinkodierendes Gen mit bisher unbekannter Funktion, zeigte eine signifikante Hypermethylierung muskulärem Herzgewebe im Vergleich zu Gefäßstrukturen, wobei die Hälfte der DM-CpG-Loci (5/10 CpGs) in Assoziation mit DHS-Regionen, d.h. potenziellen TF-Bindestellen (TFBSs), vorlagen. Die DNA-Methylierung kann die Affinität von TFs zu TFBSs beeinflussen [168] und die Bindung von TFs an TFBSs verhindern [168 - 171]. Zudem zeigen TFBSs häufig geringe DNA-Methylierungslevel [491 - 493] und die TF-TFBS-Interaktion konnte in Verbindung mit einer fehlenden DNA-Methylierung nachgewiesen werden [494, 495]. Eine Hypermethylierung der potenziellen TFBSs in *DIP2C*

könnte somit möglicherweise auf eine transkriptionelle Inaktivität dieses Gens in muskulärem Herzgewebe hindeuten.

PRDM16, welches für ein TF kodiert, der u.a. in der Determination und Differenzierung brauner Fettzellen involviert ist [496, 497], zeigte eine signifikante Hypermethylierung in Herzmuskelgewebe im Vergleich zu Gefäßstrukturen. Zwar konnte kein Zusammenhang der DM-CpG-Loci mit regulatorischen Regionen festgestellt werden, jedoch lagen alle DM-CpG-Loci innerhalb des Genkörpers von *PRDM16*. Die DNA-Methylierung in Exons stellt einen wesentlichen Auslöser von C→T-Transitions-Mutationen dar (durch Desaminierung von 5-Methyl-Cytosin zu Uracil) [498], welche eine häufige Ursache genetischer Erkrankungen [499] und krebs-auslösender Mutationen in somatischen Zellen sind [500]. Somit könnte das Auftreten der Hypermethylierung im Genkörper von *PRDM16* in muskulärem Herzgewebe potenziell C→T-Transitions-Mutationen und so eine Beeinträchtigung der Proteinfunktion bewirken. Auch in der Studie von Meurs and Kuan (2011) [501] konnte gezeigt werden, dass die DNA-Methylierung in ursächlichem Zusammenhang mit einer erhöhten Mutationsrate bestimmter kardialer Gene steht. Mutationen in *PRDM16* bzw. chromosomale Aberrationen in Überlappung mit der Genregion von *PRDM16* konnten zuvor in ursächlichem Zusammenhang mit der LVNC und der DCM identifiziert werden [321]. Funktionelle Studien von Arndt et al. (2013) [321] lassen eine gestörte proliferative Kapazität während der Herzentwicklung als primären Mechanismus von *PRDM16*-assoziierten Kardiomyopathien vermuten. Zusammenfassend könnte somit eine Beeinträchtigung der Proteinfunktion in muskulärem Herzgewebe durch methylierungsbedingte C→T-Transitions-Mutationen in *PRDM16* einen pathogenen Faktor in der Myokard-Dysfunktion der DCM darstellen.

Eine differenzielle DNA-Methylierung in Herzgewebe von Patienten mit DCM ließ sich zuvor schon in der Studie von Haas et al. (2013) [178] zeigen, indem mittels Infinium HumanMethylation 27 BeadChip Assay (27K-Array, Illumina) die DNA-Methylierung von LV-Biopsaten von DCM-Fällen mit denen aus Herzexplantaten von Transplantationspatienten als Kontrollen verglichen wurden. Ein Drittel der identifizierten CpG-Loci zeigte in dieser Studie eine Hypermethylierung bei DCM-Patienten verglichen mit den Kontrollen. Letztgenannte CpG-Loci zeigten jedoch keine Überlappung mit den DM-CpG-Loci, die in der vorliegenden Arbeit bei dem DCM-Patienten detektiert wurden. Dieser Umstand könnte zum einen durch das Vorliegen unterschiedlicher Vergleichsgruppen (Vergleich verschiedener Gewebe vs. Vergleich zu Kontrollen) und zum anderen durch mögliche zusätzliche DNA-Methylierungseffekte der Kontrollgruppe der Studie von Haas et al. (2013) [178] erklärbar sein. Die DNA-Methylierungsmuster der Kontrollgruppe letzterer Studie standen potenziell unter dem Einfluss

krankheitsspezifischer Effekte durch Herzerkrankungen der Transplantationspatienten, sowie dem Einfluss von Immunosuppressiva [502]. Zudem umfasst der 27K-Array lediglich ~ 5 % der CpG-Loci des in dieser Arbeit angewendeten 450K-Arrays [286, 394], sodass ein Großteil der in dieser Arbeit detektierten CpG-Loci in der Studie von Haas et al. (2013) [178] nicht identifiziert werden konnte.

Insgesamt war auffällig, dass die meisten identifizierten CpG-Loci (85 %) in muskulärem Gewebe verglichen mit Gefäßstrukturen hypermethyliert vorlagen, wobei insbesondere CpG-Loci regulativer DNA-Regionen (Enhancer-assoziierte CpG-Loci mit OR = 3,6) betroffen waren. Zudem konnte eine Anreicherung von Genen mit aktinbindender Funktion in Assoziation mit hypermethylierten CpG-Loci im Herzmuskelgewebe verglichen mit Gefäßstrukturen beobachtet werden. Dies ließ zum einen die Feststellung zu, dass vornehmlich Gene, die in der Ausbildung des Muskelgewebes involviert sind, ein differenzielles DNA-Methylierungsprofil zeigten. Zum anderen könnte die Hypermethylierung von CpG-Loci regulativer Elemente in muskulärem Gewebe möglicherweise in einer aberranten Expression von Muskel-assoziierten Genen und somit in der Ausprägung der DCM, einer Erkrankung des Herzmuskels, resultieren.

Zusammenfassend wurde in der vorliegenden Arbeit erstmalig eine globale DNA-Methylierungsanalyse verschiedener Herzregionen eines DCM-Patienten durchgeführt und signifikant differenzielle DNA-Methylierungssignaturen im Herzmuskelgewebe verglichen mit Gefäßstrukturen identifiziert. Eine aberrante DNA-Methylierung der Gene *GRIK2*, *DIP2C* und *PRDM16* könnte hierbei potenziell im Zusammenhang mit dem Auftreten von Myokardanomalien der DCM stehen.

4.2.3.2 Altersspezifische und individuelle Effekte in der DNA-Methylierung

In der Analyse von IAS-Gewebeproben von HLHS-Patienten und Patienten mit Rechtsherzfehlbildungen ließen sich signifikante Alterseffekte der DNA-Methylierung detektieren (siehe 3.2.2.1). Eine Rangregressionsanalyse der Gewebealterdaten ergab eine vermehrte Hypomethylierung von ca. 80 % der CpG-Loci bei „jungen“ IAS-Gewebeproben (mAG ≤ drei Wochen) im Vergleich zu „alten“ Gewebeproben (mAG > 30 Wochen) (Abbildung 29 und Abbildung 30). In verschiedenen Studien konnten Veränderungen des DNA-Methylierungsmusters mit zunehmendem Alter eines Individuums detektiert werden [145 - 152], wofür wahrscheinlich progressive und kumulative Umwelteinflüsse eine Rolle spielen. In der vorliegenden Arbeit konnten 11 Gene mit multiplen (≥ 4) DM-CpG-Loci identifiziert werden, von

denen das Gen *ESR1* zuvor schon in der Literatur mit einer altersabhängigen differenziellen DNA-Methylierung beschrieben wurde [146, 149]. Beide Studien konnten einen Zugewinn der DNA-Methylierung mit höherem Alter in *ESR1* detektieren, was auch in der vorliegenden Arbeit bei 4/5 DM-CpG-Loci in *ESR1* nachzuweisen war. Um Effekte durch eine altersbedingte differenzielle DNA-Methylierung zu vermeiden, wurden daher diese 11 Gene aus den nachfolgenden Teilprojekt-Analysen ausgeschlossen. Einen weiteren Einflussfaktor der detektierten differenziellen DNA-Methylierung unterschiedlicher Alterskategorien könnten postnatale dynamische Veränderungen des kardialen Methyloms darstellen. Diese konnten in der Studie von Sim et al. (2014) [319] in Mausembryonen innerhalb der frühen neonatalen Phase (erste zwei Wochen postnatal) nachgewiesen werden. In dieser neonatalen Phase treten wesentliche Veränderungen des Genexpressionsprofils des Herzens mit Verlust der proliferativen Kapazität und des regenerativen Potenzials auf [337 – 339, 386]. Eine differenzielle DNA-Methylierung „junger“ IAS-Gewebeproben (mAG ≤ drei Wochen) der frühen neonatalen Phase verglichen mit „alten“ Gewebeproben von > 30 Wochen könnte somit einen Effekt der postnatalen kardialen Remodellierung darstellen.

Altersabhängige Veränderungen der DNA-Methylierung können je nach Individuum unterschiedlich ausgeprägt sein [503]. Umweltfaktoren, Diäten und psychosoziale Faktoren können hierbei in der inter-individuellen differenziellen DNA-Methylierung eine Rolle spielen [504 - 507]. Eine solche individuelle differenzielle DNA-Methylierung ließ sich bei den IAS-Gewebeproben HLHS_12, HLHS_29 und HLHS_116 detektieren, indem diese ein teilweise abweichendes Methylierungsmuster im Kontext der jeweiligen Alterskategorie zeigten (Abbildung 30). Zusätzlich zu DNA-Methylierungseffekten durch das Alter könnten somit auch individuelle Effekte zu einem differenziellen DNA-Methylierungsmuster beigetragen haben. Auch die Herz-Mapping-Analyse zeigte individuelle Effekte, indem die atrialen Proben, DCM_175_RA, AVkompl_210_RA und TA_629, in der unsupervidierten Analyse außerhalb der Atrium-Gruppierung segregierten (siehe 3.2.2.2 Abbildung 32). Inter-individuelle Effekte der DNA-Methylierung konnten zuvor in verschiedenen Studien gezeigt werden [508 – 510], wobei auch individuelle DNA-Methylierungsunterschiede bei monozygoten Zwillingen nachgewiesen werden konnten [511]. Um die Intra-Gruppen-Heterogenität zu vermindern und so individuelle Effekte als Einflussfaktor zu vermeiden, wurden in der vorliegenden Arbeit bei Zwei-Gruppen-Vergleichen Proben mit starken individuellen Effekten anhand der CpG-Gesamtkorrelations-Analyse aus der Analyse ausgeschlossen.

4.2.3.3 Potenzielle Kandidatengene mit differenzieller DNA-Methylierung im Herzgewebe von HLHS-Patienten im Vergleich zu fetalem Herzkontrollgewebe

Mithilfe von fetalem Herzgewebe gesunder Individuen (BioCat, (Heidelberg, Deutschland)) wurde in dieser Arbeit ein Vergleich der DNA-Methylierung von Herzgewebeproben eines HLHS-Patientenkollektivs zu gesunden Normalkontrollen ermöglicht (siehe 3.2.2.5). Dies diente der Identifizierung potenziell krankheitsrelevanter aberranter DNA-Methylierungsmuster beim HLHS. Effekte durch altersbedingte DNA-Methylierungsunterschiede wurden hierbei durch den Vergleich zu „jungen“ (≤ 3 Wochen) Gewebeproben des HLHS-Probenkollektivs minimiert. Dennoch könnten weiterhin Alterseffekte auftreten, indem nach der Studie von Sim et al. (2014) [319] in der neonatalen Periode Veränderungen des kardialen Methyloms auftreten (siehe 4.2.3.2). Diese Effekte konnten im Mausmodell innerhalb der ersten zwei Wochen postnatal festgestellt werden, wobei die humane neonatale Periode bis Tag 28 nach der Geburt definiert ist [512]. In der Analyse „HLHS-Fetal“ wurden jedoch hauptsächlich Gewebeproben der sehr frühen neonatalen Periode mit 12/17 IAS-Gewebeproben (Median Alter: 4 Tage) verwendet. Lediglich eine Probe, HLHS_8, hatte den Maximal-Wert der „jungen“ Gewebekategorie mit 21 Tagen, wobei sich das Methylierungsmuster dieser Probe nicht von dem der anderen IAS-Proben unterschied. Zudem zeigten die insgesamt 17 IAS-Gewebeproben ein homogenes DNA-Methylierungsmuster mit r-Werten von 0,98 - 0,99 (siehe 3.2.2.5.1 Abbildung 62). Des Weiteren ließen sich in der Analyse zu altersbedingten DNA-Methylierungsunterschieden (siehe 3.2.2.1) größere Veränderungen des DNA-Methylierungsmusters im IAS-Gewebe erst ab einem Alter von drei Wochen feststellen, sodass die nach Sim et al. (2014) [319] beschriebenen Veränderungen des kardialen Methyloms in den analysierten humanen IAS-Gewebeproben möglicherweise später auftreten. Zudem wurden die signifikant differenziell methylierten CpG-Loci der HLHS-Fetal-Analyse mit DM-CpG-Loci der in Abschnitt 3.2.2.1 ermittelten Alterseffekte abgeglichen, wobei keine Überlappung vorlag. Altersbedingte DNA-Methylierungseffekte konnten somit in dieser Analyse nahezu ausgeschlossen werden. Die Gene *PDXDC1*, *NR2F1*, *NR2F1-AS1* und *XYLT1* (Diskussion von *XYLT1* erfolgt in Abschnitt 4.2.4) zeigten multiple DM-CpGs (≥ 2) mit DNA-Methylierungsdifferenzen von $\Delta\beta > 50\%$ im Herzgewebe von HLHS-Patienten verglichen mit dem Kontrollkollektiv fetaler Herzgewebeproben und werden im Folgenden diskutiert.

PDXDC1 kodiert für ein Pyridoxal-Dependent Decarboxylase Domain Containing 1-Protein mit bisher unbekannter Funktion und zeigte im Herzgewebe von HLHS-Patienten im Vergleich zu

fetalem Herzkontrollgewebe eine Hypermethylierung von CpG-Loci. Zwar war diese im wT-Test knapp nicht signifikant, jedoch konnte in diesem Gen bei demselben Patienten zusätzlich eine seltene Missense-Mutation, die nach Polyphen-2 als *probably damaging* eingestuft wird, identifiziert werden (siehe 3.3.1 Abbildung 53). Somit lagen sowohl auf epigenetischer als auch auf genetischer Ebene Aberrationen im Herzgewebe dieses HLHS-Patienten vor, die gemeinsam in einer beeinträchtigten Proteinfunktion resultieren und somit einen Einflussfaktor in der Pathogenese dieser Herzfehlbildung darstellen könnten.

Außerdem konnte eine signifikante Hypermethylierung von CpG-Loci der CDS von *NR2F1* beim HLHS identifiziert werden. Eine Hypermethylierung im Herzgewebe von HLHS-Patienten verglichen mit fetalem Herzkontrollgewebe bestand auch bei *NR2F1-AS1*, welches für dessen *Antisense*-RNA kodiert. *NR2F1* gehört zu den nukleären Rezeptoren, die ligandenabhängige Regulatoren mit der Funktion von Aktivatoren, Repressoren und Silencern der Transkription darstellen [419, 513]. *NR2F1* (NP_005645) zeigt auf Proteinlevel eine hohe Sequenzhomologie von 84,3 % identischer Sequenz (Pairwise Sequence Alignment EMBOSS Needle, siehe 2.1.3.2) zu *NR2F2* (NP_066285). In *NR2F2* konnte in dieser Arbeit im Rahmen der Studie von Al Turki et al. (2014) [303] eine kausale *de novo* Splice-Donor-Mutation bei einem HLHS-Fall nachgewiesen werden (siehe 3.1.2.1). In Anbetracht der hohen Sequenzhomologie zu *NR2F2* könnte eine aberrante DNA-Methylierung in der proteinkodierenden Sequenz des TF-Gens *NR2F1* sowie auch von *NR2F1-AS1* möglicherweise in ursächlichem Zusammenhang mit der Pathogenese des HLHS stehen. Da die DNA-Methylierung von exonischen Sequenzen potenziell C→T-Transitions-Mutationen auslösen kann [498] (in 4.2.3.1.2 diskutiert), könnte durch die Hypermethylierung die Funktion der TF-Gene *NR2F1* und *NR2F1-AS1* beeinträchtigt werden und somit einen Einflussfaktor für die Entstehung der Herzfehlbildung darstellen. Eine aberrante DNA-Methylierung von TF-Genen als pathogener Mechanismus in der Entstehung von CHDs konnte auch in der Studie von Sheng et al. (2013) [179] identifiziert werden, indem hier eine aberrante DNA-Methylierung von *NKX2.5* und *HAND1* in Assoziation zu einer negativ korrelierten mRNA-Expression dieser TF-Gene in kardialen Gewebe von Patienten mit TOF gezeigt werden konnte.

Eine vergleichende Analyse des DNA-Methylierungsmusters von primärem Herzgewebe von HLHS-Patienten und der DNA-Methylierung von Herzkontrollgewebe gesunder Individuen wurde bisher in der Literatur noch nicht beschrieben und damit in der vorliegenden Arbeit erstmalig vorgenommen. Erst kürzlich wurde in der Studie von Serra-Juhé et al. (2015) [514] eine erste Analyse des DNA-Methylierungsprofils von Patienten mit CHDs durchgeführt. In dieser Studie erfolgte allerdings ein Vergleich der DNA-Methylierung von fetalem Herzgewebe von CHD-

Patienten mit einem Kontrolldatensatz von Blutproben adulter Individuen (Alter 19 – 101 Jahre). Daher könnten für die detektierte differenzielle DNA-Methylierung dieser Studie sowohl gewebespezifische Methylierungseffekte [515] (Blut vs. Herzgewebe) als auch Alterseffekte [145 - 152] (fetal vs. 19 - 101 Jahre) verantwortlich sein.

Interessanterweise zeigten die identifizierten CpG-Loci des Vergleichs der DNA-Methylierung vom Herzgewebe von HLHS-Patienten und fetalem Herzkontrollgewebe eine hoch signifikante Anreicherung regulatoriver DNA-Regionen (RDMR- und Enhancer-assoziierte CpG-Loci, je $OR > 5,5$ mit $p \leq 4,3 \times 10^{-7}$). Zudem ließen sich stark erhöhte DNA-Methylierungsdifferenzen von $\Delta\beta > 45\%$ im Herzgewebe von HLHS-Patienten verglichen mit fetalem Herzkontrollgewebe detektieren. In Anbetracht des Befundes, dass eine signifikante Anreicherung von Entwicklungsgenen innerhalb der detektierten DM-CpG-Loci vorlag, könnte eine aberrante DNA-Methylierung von vorwiegend regulatoriven Elementen in Entwicklungsgenen in einer abnormen Herzentwicklung wie beim HLHS resultieren.

4.2.4 Zusammenhänge auf Genom-, Epigenom- und

Transkriptomebene in der Entstehung des HLHS

In dieser Arbeit wurden genetische Analysen auf Keimbahn- und somatischer Ebene, sowie DNA-Methylierungsanalysen und Transkriptomanalysen durchgeführt, die abschließend in einer integrativen Analyse auf Überlappungen betroffener Gene und *Gen-Pathways* auf Genom-, Epigenom- und Transkriptomebene untersucht wurden. Kombinationen von Ereignissen auf diesen verschiedenen Ebenen könnten im besonderen Maße darauf hinweisen, dass ursächliche Zusammenhänge in der Pathogenese des HLHS bestehen. Die Ergebnisse der integrativen Analysen (siehe 3.3) werden im Folgenden diskutiert.

Gene mit einer differenziellen DNA-Methylierung und funktionellen Genvarianten bei Patienten mit HLHS

Es konnten insgesamt 20 Gene mit seltenen bzw. bisher unbekanntem proteinschädigenden Mutationen bei Patienten mit HLHS nachgewiesen werden, die zusätzlich im Herzgewebe von Patienten mit unterschiedlichen CHDs, inklusive dem HLHS, differenzielle DNA-

Methylierungsmuster aufwiesen. Diese umfassten u.a. das TF-kodierende Gen *NR2F2*, sowie Gene, die in der Ausbildung kardialer Strukturen involviert sind (*DOCK1* und *CTNNB1*).

Wie bereits in Abschnitt 4.2.1 diskutiert, konnte in *NR2F2* eine kausale *de novo* Splice-Donor-Mutation bei einem Patienten mit HLHS im Rahmen der Studie von Al Turki et al. (2014) [303] nachgewiesen werden. Zusätzlich ließ sich in diesem Gen eine differenzielle Methylierung multipler CpG-Loci in verschiedenen Herzregionen von CHD-Patienten identifizieren (siehe 3.3.1 Abbildung 51). Die DM-CpG-Loci zeigten hierbei eine Hypomethylierung in ventrikulärem, atrialem und fetalem Herzgewebe, während das Herzklappengewebe und Gefäß-Strukturen hypermethyliert vorlagen. Die Hypomethylierung in atrialem und ventrikulärem Gewebe in *NR2F2* könnte mit einem tendenziell weniger differenzierten Gewebestatus assoziiert sein, da auch die fetalen Gewebeproben der 25. und 40. SSW eine ähnliche DNA-Methylierung dieser Loci aufwiesen. Die höhere DNA-Methylierung in Klappen- und Gefäßgewebe von *NR2F2* könnte hierbei in Zusammenhang mit dessen wahrscheinlicher Funktion in der endothelial-mesenchymalen Signalgebung der Herzentwicklung stehen [420]. Da deren Signalgebung im Prozess der EndMT zur Herzklappenentwicklung involviert ist, könnte eine differenzielle DNA-Methylierung von *NR2F2* einen epigenetischen Faktor in Differenzierungsprozessen der Herzklappen- und Gefäßentwicklung darstellen. Potenziell könnte eine aberrante DNA-Methylierung von CpG-Loci in *NR2F2* in der pathologischen Herzklappenentwicklung bei LVOTOs wie dem HLHS involviert sein.

Zusätzlich zur Hypermethylierung eines CpG-Locus in muskulärem Herzgewebe verglichen mit Gefäßstrukturen konnten in *DOCK1* zwei bisher unbekannte Missense-Mutationen, als *probably* bzw. *possibly damaging* eingestuft, detektiert werden. Diese lagen beim Patienten HLHS_132 nach Sanger-Sequenzierungsanalyse des Eltern-Indexpatient-Trios *compound* heterozygot vor (siehe 3.3.1 Abbildung 52). Während eine heterozygote Ausprägung keinen Phänotyp bei den gesunden Eltern auslöst, könnten die prädiktiv proteinschädigenden Mutationen beider Allele im *compound* heterozygoten Zustand beim Indexpatienten HLHS_132 in einem Funktionsverlust dieses Gens resultieren und somit einen krankheitsrelevanten Faktor für die Entstehung des HLHS darstellen. *DOCK1*, welches für einen Guanin-Nukleotid-Austauschfaktor kodiert, der als Aktivator der GTPase Rac1 fungiert [516, 517], ist in der Myoblasten-Fusion [518] und in der kardiovaskulären Entwicklung involviert [519]. *DOCK1* kontrolliert hierbei die endotheliale Zellmigration während der Kardiogenese [519] und schützt Endothelzellen durch die Aktivierung des Rac1/PAK/AKT (Rac1/p21-aktivierte Kinase/Proteinkinase B)-Signalweges vor Apoptose. *DOCK1* wirkt somit als stabilisierender Faktor des Endothels [520]. Ein Knockout im Mausmodell von *DOCK1* resultierte in kardialen Defekten, u.a. VSDs, sowie in Verdickungen und Fusionen

der Mitralklappensegel [518, 519]. Das *compound* heterozygote Auftreten zweier bisher unbekannter proteinschädigender Mutationen sowie kardiale Defekte im Mausmodell, die u.a. die Mitralklappe – welche auch beim HLHS Anomalien aufzeigt – betreffen, lassen vermuten, dass das *DOCK1*-Gen ein wahrscheinlich krankheitsrelevantes Kandidatengen in der Entstehung des HLHS darstellt.

In *CTNNB1* konnten in dieser Arbeit eine seltene, als *probably damaging* eingestufte Missense-Mutation, sowie eine Hypermethylierung im IAS-Gewebe desselben HLHS-Patienten im Vergleich zu fetalem Herzkontrollgewebe detektiert werden (siehe 3.3.1 Abbildung 53). Das Genprodukt von *CTNNB1* ist der Wnt-Signalweg-Effektor β -Catenin, welcher die Transkription von Genen reguliert, die in der Proliferation und Differenzierung von Zellen in Entwicklungsprozessen involviert sind [521]. Der *Ctnnb1*-abhängige Wnt-Signalweg ist in multiplen Phasen der kardialen Entwicklung von wesentlicher Bedeutung. Zum einen ist dieser Signalweg in der Mesodermentstehung während der Gastrulation involviert, einer Grundvoraussetzung für die kardiale Entwicklung [522]. Zum anderen ist der Wnt- β -Catenin-Signalweg für die Aufrechterhaltung der SHF-Proliferationsfähigkeit von Bedeutung [21]. Eine mesenchymale Inaktivierung von *CTNNB1* in Mäusen zeigte nach der Studie von Huh and Ornitz (2010) [523] ein DiGeorge-Syndrom⁴⁷-ähnliches phänotypisches Spektrum inklusive kardialer Defekte, wie u.a. eine Hypoplasie von Pulmonalarterie und Aorta sowie Septationsdefekte. Kombinatorische Effekte aus einer aberranten DNA-Methylierung sowie proteinschädigenden Mutationen in *CTNNB1* könnten somit einen möglichen Einflussfaktor in der Entstehung der Hypoplasie aortaler Strukturen beim HLHS darstellen.

***ACE2* und *SOSTDC1* mit einer differenziellen Genexpression und funktionellen Genvarianten als mögliche Kandidatengene in der Entstehung des HLHS**

Die Gene *ACE2* und *SOSTDC1* konnten in dieser Arbeit als mögliche Kandidatengene in der Entstehung des HLHS identifiziert werden, da diese Gene, welche zuvor auch schon in der Literatur in Assoziation zu CHDs beschrieben wurden, sowohl eine aberrante Genexpression als auch proteinschädigende Mutationen im IAS-Gewebe von HLHS-Patienten aufwiesen.

In *ACE2* konnte eine bisher unbekannte hemizygote Missense-Mutation (als *probably damaging* eingestuft) bei HLHS_4 identifiziert werden und dieses Gen zeigte bei Analyse desselben IAS-Gewebes dieses HLHS-Patienten eine signifikante Hochregulation verglichen mit Patienten mit

⁴⁷ DiGeorge-Syndrom umfasst ein angeborenes Fehlbildungssyndrom, gekennzeichnet durch eine Hypoplasie des Thymus und der Nebenschilddrüsen, einer Gesichtsdysmorphie und CHDs (OMIM 188400).

Rechtsherzfehlbildungen (siehe 3.3.2 Abbildung 54). Auch in der Studie von Goulter et al. (2004) [524] konnte eine gesteigerte Expression von *ACE2* in Herzgewebeproben, dem ventrikulärem Myokard, von Patienten mit Herzfehlbildungen (Kardiomyopathie) detektiert werden, sodass eine hochregulierte Genexpression von *ACE2* in Assoziation mit der Entstehung von kardialen Anomalien steht. *ACE2* (Angiotensin-konvertierendes Enzym) ist ein wesentlicher Regulator des Renin-Angiotensin-Systems [525, 526]. Bei Deletionen von *ACE2* konnten im Mausmodell kardiale Kontraktilitätsdefekte gezeigt werden [526], sodass *ACE2* einen wichtigen Regulator der kardialen Funktion darstellt. Eine Inhibition von Angiotensin-konvertierenden Enzymen sowie die Blockierung von Angiotensin II-Rezeptoren zeigte zudem ein klinisch signifikant verbessertes Behandlungsergebnis bei Patienten mit Herzinsuffizienz [527], sodass eine erhöhte Genexpression von *ACE2* auch in einem ursächlichen Zusammenhang mit der Entstehung von Anomalien des Herzens stehen könnte. Die bei HLHS_4 hemizygot ausgeprägte Missense-Mutation 2188A>C (p.T730P) könnte zudem eine aktivierende Mutation darstellen, da die Transition von Tryptophan zu Prolin, das besonders viele *cis*-Amidbindungen aufweist [528], veränderte Eigenschaften der Proteinfaltung beinhalten kann. In der phänotypisch gesunden Mutter lag diese Mutation heterozygot vor, sodass möglicherweise beide Allele für eine pathogene Wirkung betroffen sein müssen.

In *SOSTDC1* konnte im IAS-Gewebe von HLHS-Patienten im Vergleich zu Patienten mit Rechtsherzfehlbildungen eine herunterregulierte Genexpression detektiert werden. Darüber hinaus ließ sich eine LOF-Mutation (Frameshift), als *probably damaging* eingestuft, bei HLHS_132 nachweisen (siehe 3.3.2 Abbildung 55). Ein im Rahmen dieser RNA-Sequenzierungsanalysen durchgeführter Knockdown im Zebrafisch-Modell ergab u.a. eine kardiale Dysfunktion, d.h. ein kardiales Ödem. Ein Funktionsverlust dieses Gens zeigte somit eine Assoziation zu einem kardialen Defekt (noch nicht publiziert). *SOSTDC1* wurde zuvor als Wnt-Antagonist [529], sowie als Inhibitor des BMP-Signalweges beschrieben [530 - 532]. Beide Signalwege haben eine wichtige Funktion in der Kardiogenese. Die Wnt-Signalgebung ist hierbei in der Expansion kardialer Vorläuferzellen von wesentlicher Bedeutung [533 - 535], und deren nachfolgende Differenzierung wird durch eine BMP-Signalgebung und Wnt-Inhibition reguliert [536, 537]. Somit könnte eine aberrante Genexpression in *SOSTDC1* einen Einflussfaktor für die Entstehung von Herzfehlbildungen darstellen. Zwar ließ sich die LOF-Mutation in *SOSTDC1* des HLHS-Falls HLHS_132 ebenfalls beim phänotypisch gesunden Vater detektieren, jedoch könnte ein kumulatives Auftreten von Mutationen in Genen der kardialen Entwicklung, wie z.B. die bei HLHS_132 zusätzlich detektierte *compound* heterozygote Mutation in *DOCK1*, in der Pathogenese des HLHS eine Rolle spielen.

Mögliche pathologische Veränderungen des Proteoglykanmetabolismus durch differenzielle DNA-Methylierung und dysregulierte Genexpression von *XYLT1* bei Patienten mit HLHS

Das Gen *XYLT1* (Xylosyltransferase 1), das in der Synthese von Glykosaminoglykan-enthaltenden Proteoglykanen involviert ist [538 - 540], zeigte im Herzgewebe von HLHS-Patienten verglichen mit fetalem Herzkontrollgewebe eine signifikante Hypomethylierung zweier Enhancer-assoziiierter CpG-Loci und eines DHS-assoziierten CpG-Locus. Eine verringerte Methylierung dieser Enhancer-Regionen könnte in einer erhöhten transkriptionellen Aktivität resultieren, indem eine Bindung von Aktivatoren aufgrund der fehlenden DNA-Methylierung ermöglicht wird [176, 483].

Im IAS-Gewebe dieser HLHS-Patienten konnte zudem eine signifikante Hochregulation der Genexpression ($p = 0,002865$) im Vergleich zu Patienten mit Rechtsherzfehlbildungen detektiert werden (siehe 3.3.3 Abbildung 56). Dies würde somit auf Transkriptomebene den epigenetischen Mechanismus der Enhancer-Hypomethylierung mit resultierender Erhöhung der transkriptionellen Aktivität bestätigen. Verschiedene Studien konnten zeigen, dass eine erhöhte Xylosyltransferase-Aktivität in Zusammenhang mit fibrotischen und sklerotischen Veränderungen des Bindegewebes auftritt [541 - 543]. Mutationen in *XYLT1* konnten bei Fällen mit Pseudoxanthoma *elasticum*, einer vererbaren degenerativen Bindegewebserkrankung, identifiziert werden [543]. Eine aberrante DNA-Methylierung in *XYLT1* in Zusammenhang mit einer hochregulierten Genexpression könnte somit einen möglichen Einflussfaktor in der Pathogenese des HLHS darstellen. Die dysregulierte *XYLT1*-Expression könnte hierbei in Veränderungen des Proteoglykanmetabolismus resultieren, sodass die Ausbildung bindegewebshaltiger Strukturen des Herzens gestört ist. Dies könnte in direktem Zusammenhang mit dem Auftreten von einer Endokardfibroelastose (EFE) des LVs stehen, die häufig beim HLHS bei Durchgängigkeit der Mitralklappe beobachtet werden kann [544, 545]. Diese abnormale Verdickung des bindegewebshaltigen Endokards [546, 547] spielt in der Progression der linksventrikulären Wachstumsverzögerung sowie der diastolischen Dysfunktion des fetalen Herzens beim HLHS eine Rolle [548, 549]. Somit könnten die in dieser Arbeit nachgewiesene Aberration in der DNA-Methylierung sowie die differenzielle Genexpression von *XYLT1* durch konsekutiv pathologische Veränderungen des Proteoglykanmetabolismus in der Progression der LV-Dysfunktion beim HLHS involviert sein.

Insgesamt ließen sich in der vorliegenden Arbeit verschiedene Gene der transkriptionellen Regulation (*NR2F2* und *ACE2*), der Ausbildung kardialer Strukturen (*DOCK1* und *XYLT1*) sowie der Signaltransduktion (*CTNNB1* und *SOSTDC1*) in einem **potenziell ursächlichen Zusammenhang mit der Pathogenese des HLHS** identifizieren. Kumulative Ereignisse auf genetischer, epigenetischer und transkriptioneller Ebene könnten hierbei einen möglichen molekularen Mechanismus in der Entstehung des HLHS darstellen. Die Entwicklung einer solchen komplexen Herzfehlbildung könnte auf kombinatorischen Effekten entsprechend eines „*Two Hit*“-Modells basieren, welches als möglicher Pathomechanismus bei Entwicklungsstörungen (*developmental disorder* (DD)) diskutiert wurde [550, 551]. In der Studie von Girirajan et al. (2010) [550] konnte beobachtet werden, dass DD-Patienten mit einer (teilweise vererbten) Mikrodeletion in 16p12.1 mit einer höheren Wahrscheinlichkeit zusätzliche CNVs („*Two Hit*“) aufzeigten als die Kontrollkohorte. Dieser zweite „*Hit*“ könnte nach Girirajan et al. (2010) den Schweregrad des klinischen Spektrums, welches u.a. auch Patienten mit HLHS umfasste, beeinflussen. Diese Daten unterstützen ein oligogenes Vererbungsmodell, welches als wahrscheinlicher Vererbungsmechanismus bei CHDs diskutiert wird [78, 236, 246, 304]. Kürzlich konnte in kausalem Zusammenhang mit nicht-syndromalen CHDs wie dem HLHS eine signifikante Anreicherung von LOF-Mutationen, vererbt durch nicht-betroffene Eltern, gezeigt werden (Sifrim und Hitz et al. 2016⁴⁸, eigene Mitwirkung). Somit könnten vererbte Genvarianten als prädisponierende Faktoren in Kombination mit einem zweiten „*Hit*“, genetisch oder epigenetisch, ein mögliches Modell der Entstehung von CHDs wie dem HLHS darstellen [465, 552].

Hypothetisch könnte beim HLHS das Auftreten kumulativer Genvarianten „primäre“ anatomische Defekte linksseitiger kardialer Strukturen auslösen, deren konsekutiv veränderte Blutflussdynamik sekundär eine pathologische Morphogenese des Herzens durch epigenetische Veränderungen von Genexpressionsprofilen bewirkt [205, 206]. *In vitro*-Studien konnten bei Fluss-induziert wirkenden Kräften eine Veränderung der Zytoskelettstruktur und des Genexpressionsprofils in kultivierten kardialen Endothelzellen detektieren, sodass die Flusssdynamik eine wichtige Funktion in der kardialen Morphogenese haben könnte [553, 554]. Der Fortschritt in der Ultraschall-Technik ließ zudem erkennen, dass die Hypoplasie des linken Ventrikels während der Schwangerschaft zunimmt [189]. Da die Proliferationsfähigkeit des Herzens zwar noch in der frühen Embryogenese besteht [555, 556], es jedoch in späteren Entwicklungsstadien zu einem Verlust der proliferativen Kapazität und des regenerativen Potenzials kommt [337 - 339], ist die Hypoplasie des linken Ventrikels nicht reversibel.

⁴⁸ Dokument derzeit in Revision, eingereicht bei Nature Genetics (siehe Abschnitt 11)

Zusammenfassend könnte das Auftreten kombinatorischer Ereignisse auf genetischer und epigenetischer Ebene in Genen von Entwicklungs-*Pathways*, der transkriptionellen Regulation und der Kardiogenese, wie dies in der vorliegenden Arbeit detektiert wurde, einen möglichen Pathomechanismus in der Entstehung komplexer CHDs wie dem HLHS darstellen.

4.3 Fazit

Das HLHS stellt eine der schwerwiegendsten menschlichen Fehlbildungen mit hoher Mortalität und Langzeit-Morbidität dar [182 – 184]. Der zugrundeliegende Pathomechanismus ist bis heute noch weitgehend unklar [188]. Eine variable Penetranz detektierter Genvarianten [200, 218 - 220] und eine unterschiedliche Ausprägung linksseitiger kardialer Defekte bei Familien mit HLHS [233, 237, 246] lassen vermuten, dass neben einer vererbten genetischen Komponente weitere Faktoren in der Ätiologie des HLHS involviert sind. Die Ergebnisse der vorliegenden Arbeit unterstützen diese Theorie. Einerseits ließen sich auf Keimbahnebene potenziell krankheitsrelevante und kausale Genvarianten nachweisen und ein autosomal dominantes Vererbungsmodell mit reduzierter Penetranz konnte als wahrscheinlichstes Vererbungsmodell in einer Familie mit HLHS identifiziert werden. Andererseits konnten auf somatischer Ebene im Herzgewebe von Patienten mit HLHS neben potenziell krankheitsrelevanten Genvarianten auch eine differenzielle DNA-Methylierung und Genexpression detektiert werden. In Anbetracht des schwer zugänglichen Gewebematerials liegen bisher in der Literatur nur wenige Daten über Mutationen und epigenetische Veränderungen im Herzgewebe von Patienten mit kardialen Defekten vor. Somit kann die vorliegende Arbeit zu einem besseren Verständnis der molekularen Mechanismen des HLHS beitragen.

Darüber hinaus konnten erstmalig gewebespezifische differenzielle DNA-Methylierungsprofile in primärem Herzgewebe unterschiedlicher Herzregionen von Patienten mit CHDs, inklusive dem HLHS, gezeigt werden. Hierbei ließen sich zum einen Atrium- und Ventrikel-spezifische Methylierungsmuster, sowie eine differenzielle Muskel- und Gefäßsignatur nachweisen. Diese Erkenntnisse liefern einen Beitrag zur Charakterisierung von Herzzellen in der Stammzellforschung. Die herzregionspezifischen Methylierungsprofile können potenziell als epigenetische Marker zur Unterscheidung der Zelltypen Verwendung finden.

In einer abschließenden gesamtintegrativen Analyse konnten Gene der transkriptionellen Regulation, der Ausbildung kardialer Strukturen sowie der Signaltransduktion in einem potenziell ursächlichen Zusammenhang mit der Entstehung des HLHS identifiziert werden, da diese Gene überlappend auf Genom-, Epigenom- oder Transkriptomebene Aberrationen aufwiesen. Es besteht die Möglichkeit, dass bei einer Anhäufung solcher genetischen und epigenetischen Veränderungen die Entstehung des HLHS begünstigt wird.

Zusammenfassend liefern die Ergebnisse einen Beitrag zur Aufklärung der Pathomechanismen des HLHS. Die identifizierten Gene und *Pathways* können in weiterführenden funktionellen Studien zur Analyse einer Genotyp-Phänotyp-Korrelation und möglicherweise zur Entwicklung neuer Therapieansätze nutzbar gemacht werden.

4.4 Ausblick

In der vorliegenden Arbeit ließen sich neben neuen ursächlichen Genvarianten auf Keimbahnebene auch Assoziationen auf somatischer Ebene, die möglicherweise mit der Pathogenese des HLHS zusammenhängen, identifizieren. Es ist daher davon auszugehen, dass sowohl genetische als auch epigenetische Veränderungen in der Entstehung dieser komplexen Herzfehlbildung eine Rolle spielen. Die analysierten HLHS-Fälle zeigten zwar jeweils phänotypisch gesunde Elternteile, grundsätzlich könnten hierbei jedoch asymptomatische und damit undiagnostizierte milde Ausprägungen von AoV-Defekten bei Familienmitgliedern dieser HLHS-Fälle vorliegen. Zukünftig wären daher systematische Echokardiographie-Analysen von Familienmitgliedern dieser Fälle notwendig, um das in dieser Arbeit im Rahmen der WES-Studie zu Familien mit LVOTOs (Hitz und Hoff et al., Publikation in Vorbereitung (siehe 11)) identifizierte autosomal dominante Vererbungsmodell mit reduzierter Penetranz weiter validieren zu können. Ein weiterer möglicher Mechanismus für das Auftreten von variabler Penetranz beim HLHS könnte ein genetischer Mosaizismus in bestimmten Herzgeweberegionen sein, der zu LS-CHDs unterschiedlichen Schweregrads führt. Aufgrund der Limitationen der Sanger-Sequenzierungstechnologie ließen sich die in dieser Arbeit identifizierten herzwespezifischen Mutationen bisher noch nicht verifizieren. Es sind daher *Deep-Sequencing*-Analysen zur Identifizierung potenzieller somatischer Mosaik-Mutationen im Herzgewebe von Patienten mit HLHS geplant.

Weiterhin konnte in der vorliegenden Arbeit erstmalig in Herzgewebeproben eine differenzielle DNA-Methylierung unterschiedlicher Herzregionen gezeigt werden. Ein solches differenzielles DNA-Methylierungsprofil im Atrium- und Ventrikel-Gewebe, sowie im Herzmuskelgewebe und Gefäßstrukturen könnte potenziell als epigenetischer Marker in der Stammzellforschung Anwendung finden. Diese Informationen könnten eine effiziente *in vitro*-Differenzierung von atrialen und ventrikulären Kardiomyozyten ermöglichen und zur weiteren Validierung bzw. Unterscheidung dieser beiden Zelltypen herangezogen werden. Um solche Analysen patientenspezifisch durchzuführen, ist zukünftig die Generierung von iPSCs bei Patienten mit CHDs, insbesondere mit HLHS, in Zusammenarbeit mit dem Institut für Klinische und Experimentelle Pharmakologie und Toxikologie des Universitätsklinikums Hamburg-Eppendorf geplant. Die generierten patientenspezifischen iPSCs und die daraus *in vitro*-differenzierten Kardiomyozyten sollen in weiterführenden Studien gesamtgenomisch und funktionell charakterisiert werden, um mögliche Pathomechanismen zu untersuchen. Auch sind zur Analyse von Einflüssen auf die frühe kardiale Entwicklung funktionelle Studien von identifizierten Kandidatengen im Zebrafisch-Modell geplant. Im Rahmen der vorliegenden Arbeit erfolgte hierfür parallel der Aufbau einer Zebrafisch-Laboreinheit in der Klinik für angeborene Herzfehler und Kinderkardiologie (UKSH, Campus Kiel) und erste mutante Zebrafisch-Linien wurden generiert.

Insgesamt können die in dieser Arbeit gewonnenen Erkenntnisse über genetische, epigenetische und transkriptionelle Veränderungen bei Patienten mit HLHS einen Beitrag für ein besseres Verständnis der pathogenetischen Grundlagen des HLHS liefern. Mithilfe neuer NGS-Technologien, wie in dieser Arbeit angewendet, sowie mithilfe von Modellorganismen und patientenspezifischen iPSCs lassen sich zukünftig Fortschritte in der Risikoprädiktion von CHDs bei familiären Fällen erreichen und neue therapeutische Ansätze durch patientenspezifische Stammzelltechnologien ermöglichen.

5 Zusammenfassung

Das hypoplastische Linksherzsyndrom (HLHS) stellt ein kongenitales Herz- und Gefäßfehlbildungssyndrom dar, welches durch eine Atresie oder Stenose der Aorten- und Mitralklappe sowie eine Hypoplasie des linken Ventrikels und der Aorta gekennzeichnet ist. Mit einer Inzidenz von 2–3 % aller Herzfehlbildungen zählt das HLHS zwar zu den seltenen kongenitalen Herzfehlbildungen (CHDs), ist aber unbehandelt ursächlich für 25–40 % der neonatalen kardialen Todesfälle und gehört somit zu den schwerwiegendsten CHDs. Während sich wesentliche Fortschritte in der Diagnose und Therapie erzielen ließen, ist die Ätiologie des HLHS bis heute weitgehend unklar. Betrachtet man die Herzentwicklung, so stellt diese einen komplexen, räumlich und zeitlich fein abgestimmten Prozess dar, der auf einer koordinierten Interaktion genetischer sowie epigenetischer Faktoren basiert. Störungen dieser Mechanismen der Kardiogenese können zu einer fehlerhaften Herzentwicklung führen und somit in strukturellen Herzfehlbildungen wie dem HLHS resultieren.

Um die molekularen Mechanismen in der Entstehung des HLHS zu untersuchen, wurden in dieser Arbeit Analysen auf Genom-, Epigenom- und Transkriptomebene an Keimbahn- bzw. Herzgewebematerial von Patienten mit HLHS und anderen CHDs durchgeführt. Dies umfasste auf Keimbahnebene Gesamt-Exom-Sequenzierungen (WES)-Analysen an drei Familien mit Catel-Manzke-Syndrom, als Beispiel eines syndrom-assoziierten CHDs, sowie Sanger-Sequenzierungen von Kandidatengenen bei zwei Familien mit rekurrent auftretenden nicht-syndromalen HLHS-Fällen. Zusätzlich erfolgte eine SNP-Array-Analyse eines HLHS-Falls im familiären Kontext, um mögliche krankheitsrelevante Kopienzahlvariationen zu identifizieren. Auf somatischer Ebene wurden neun Herzgewebeproben von HLHS-Patienten in einer WES-Analyse hinsichtlich pathogener Genvarianten untersucht und Kandidatengene mittels Sanger-Sequenzierung verifiziert. Zusätzlich wurden RNA-Sequenzierungen an Herzgewebeproben von 26 HLHS-Patienten und 10 Patienten mit Rechtsherzfehlbildungen zur Analyse von Veränderungen auf Transkriptomebene untersucht. Darüber hinaus erfolgten auf epigenetischer Ebene Analysen mittels Infinium HumanMethylation450 BeadChip Array von insgesamt 78 Hybridisierungen von Herzgewebeproben unterschiedlicher CHD-Patienten, inklusive 27 Herzgewebeproben von HLHS-Patienten. Mittels Bisulfit-Pyrosequenzierung und dem massenspektrometrischen Verfahren der Epityper-Technik wurden ausgewählte Ergebnisse verifiziert. In einer gesamtintegrativen Analyse wurden abschließend die identifizierten Kandidatengene und Gen-*Pathways* hinsichtlich Überlappungen auf Genom-, Epigenom- und Transkriptomebene untersucht.

In dieser Arbeit konnten auf Keimbahnebene bei drei Familien mit Catel-Manzke-Syndrom, bei dem CHDs vorkommen, kausale Genvarianten im *TGDS*-Gen identifiziert werden. Des Weiteren ließ sich eine kausale bisher unbekannte *de novo* Splice-Donor-Mutation in *NR2F2* bei einem nicht-syndromalen Fall mit HLHS nachweisen. In einer Familie mit zwei Geschwistern mit HLHS konnten potenziell krankheitsrelevante Genvarianten in *C15orf62*, *TNKS1BP1* und *PRF1* nach autosomal dominantem Vererbungsmodell mit reduzierter Penetranz identifiziert werden. Neben diesen Ergebnissen auf Keimbahnebene konnten auch auf somatischer Ebene im Herzgewebe von Patienten mit HLHS seltene bzw. bisher unbekannte funktionelle Mutationen in Genen detektiert werden, die in kardialen Entwicklungsprozessen, sowie in der (post-)transkriptionellen Regulation involviert sind. Hervorzuheben ist hierbei die Identifizierung einer wahrscheinlich kausalen proteintrunkierenden Mutation in *NOTCH1*, einem Gen, welches bereits zuvor in verschiedenen Studien in Assoziation mit dem HLHS gefunden werden konnte. Zum anderen konnte eine wahrscheinlich krankheitsrelevante *compound* heterozygote Mutation in *DOCK1*, das in der Ausbildung kardialer Strukturen involviert ist, bei einem Patienten mit HLHS identifiziert werden. In einer gesamtintegrativen Analyse ließen sich verschiedene Gene identifizieren, die in einem potenziell ursächlichen Zusammenhang mit der Entstehung des HLHS stehen, da diese Gene überlappend auf Genom-, Epigenom- oder Transkriptomebene Veränderungen aufwiesen. Dies umfasste Gene, die in der Regulation der Transkription (*NR2F2* und *ACE2*), in der Ausbildung kardialer Strukturen (*DOCK1* und *XYLT1*), sowie in der Signaltransduktion (*NOTCH1*, *CTNNB1* und *SOSTDC1*) von Bedeutung sind. Des Weiteren konnten erstmalig in primärem Herzgewebe von Patienten mit CHDs herzregionspezifische DNA-Methylierungsprofile detektiert werden. Hierbei ließen sich Atrium- und Ventrikelspezifische Methylierungsmuster als auch eine differenzielle Herzmuskel- und Gefäßsignatur nachweisen. Zusammenfassend lassen die Ergebnisse dieser Arbeit vermuten, dass kumulative Ereignisse auf genetischer und epigenetischer Ebene in einem potenziell ursächlichen Zusammenhang mit der Entwicklung des HLHS stehen. Es besteht die Möglichkeit, dass vererbte Genvarianten als prädisponierende Faktoren in Kombination mit somatischen Mutationen und/oder epigenetischen Aberrationen im Herzgewebe die Entstehung einer solchen Herzfehlbildung begünstigen. Insgesamt können die in dieser Arbeit gewonnenen Erkenntnisse aus NGS-Analysen und globalen DNA-Methylierungsanalysen sowie zukünftige funktionelle Studien im Zebrafisch-Modell und Studien mit patientenspezifischen induzierbaren pluripotenten Stammzellen einen Beitrag für ein besseres Verständnis der Ätiologie des HLHS liefern.

6 Summary

Hypoplastic left heart syndrome (HLHS) displays a severe congenital heart defect (CHD) that encompasses a spectrum of structural defects characterized by atresia or stenosis of the aortic and mitral valve, as well as a hypoplasia of the left ventricle and aorta. Despite its low incidence rate, accounting for 2 to 3 % of all CHDs, HLHS, if left untreated, is responsible for 25 to 40 % of all neonatal cardiac deaths and represents one of the most severe cardiac malformations. While substantial progress has been made in diagnosis and surgical treatment, the etiology of HLHS remains unknown. The formation of the heart is a complex spatiotemporally regulated process involving a coordinated interplay of genetic and epigenetic factors. Any perturbation of this sensitive process can lead to abnormal cardiogenesis resulting in structural cardiac defects, such as HLHS.

In order to investigate the molecular mechanisms of HLHS, heart tissue and blood samples from different patients with HLHS, as well as from patients displaying other CHDs, have been examined. Genomic, transcriptomic and epigenomic datasets have been generated and analyzed in this thesis. This included a whole exome sequencing (WES) dataset from three families with Catel-Manzke syndrome and two families with recurrent form of non-syndromic HLHS. In addition, a parent-offspring trio of an isolated HLHS case has been examined for rare disease-associated copy number variations. In order to analyze somatic changes potentially involved in the pathogenesis of HLHS, WES was performed on nine heart tissue samples from independent cases. RNA sequencing on heart tissue samples of 26 HLHS cases and of 10 cases with right-sided obstructive lesions has been carried out to identify transcriptional changes. To further evaluate epigenetic alterations, 78 heart tissue samples from patients with CHD, including 27 heart tissue samples from cases with HLHS, have been analyzed using Infinium HumanMethylation450 BeadChip Array. Downstream verification of the identified differentially methylated loci was done by applying bisulfite pyrosequencing and mass spectrometric EpiTyper technique. To summarize these findings and identify possible pathways involved in HLHS disease pathogenesis, an integrative analysis has been performed.

In this thesis, disease-causing mutations in *TGDS* have been identified in three families with Catel-Manzke syndrome, which often is accompanied by a structural heart defect. Furthermore, a *de novo* donor splice site mutation has been detected as a disease-causing variant in a case with HLHS. Additionally, in a family with two siblings with HLHS, possibly disease-causing mutations have been identified in *C15orf62*, *TNKS1BP1* und *PRF1* displaying an autosomal dominant inheritance pattern with reduced penetrance. Despite these findings, the analysis on

heart tissue samples revealed rare functional variants in genes which are possibly linked to heart development and (post-)transcriptional regulation. Notably, a protein truncating mutation in *NOTCH1* displaying a probably disease-causing variant has been identified. A probably disease-causing compound heterozygous mutation in *DOCK1*, a gene which is known to play a role in cardiac development, has been identified in a case with HLHS. Overall, the integrative approach revealed candidate genes necessary for transcriptional regulation (*NR2F2* und *ACE2*), cardiac development (*DOCK1* und *XYLT1*) and signal transduction (*NOTCH1*, *CTNNB1* und *SOSTDC1*). One of the most interesting findings of this thesis was the detection of heart region specific DNA methylation profiles. This included specific DNA methylation profiles for the atrium and ventricle, as well as differential methylation signatures of cardiac muscle and vessels.

In summary, the results of the present study suggest that the occurrence of cumulative events on genetic and epigenetic level could represent a possible mechanism in the pathogenesis of complex CHDs, such as HLHS. A possible explanation would be that inherited genetic variants, as predisposing factors, acting in combination with somatic mutations and/or epigenetic aberrations in cardiac tissue could give rise to HLHS. Overall, insights gained from NGS analyses and global DNA methylation analysis in this thesis, as well as future functional studies in zebrafish and patient-specific inducible pluripotent stem cells, could provide a contribution for a better understanding of the etiology of HLHS.

7 Quellenverzeichnis

- [1] Galen, T. A. J. B. (1916). On the Natural Faculties.
- [2] Masic, I. (2010). On occasion of 800th anniversary of birth of Ibn al-Nafis--discoverer of cardiac and pulmonary circulation. *Med. Arh.* 64, 309–313.
- [3] Harvey, W. (1628). *Exercitatio Anatomica de Motu Cordis et Sanguinis in Animalibus*.
- [4] Aird, W.C. (2011). Discovery of the cardiovascular system: from Galen to William Harvey. *J. Thromb. Haemost.* 9, 118–129.
- [5] Netter, F.H. (2010). *Atlas of Human Anatomy* (Elsevier Health Sciences).
- [6] Hinrichsen, K.V. (2013). *Humanembryologie: Lehrbuch und Atlas der vorgeburtlichen Entwicklung des Menschen* (Springer-Verlag).
- [7] Katz, A.M. (2010). *Physiology of the Heart* (Lippincott Williams & Wilkins).
- [8] Lilly, L.S. (2012). *Braunwald's Heart Disease: A Textbook of Cardiovascular Medicine* (Elsevier Health Sciences).
- [9] Heiden, K. (2009). *Congenital Heart Defects, Simplified* (Midwest EchoSolutions).
- [10] Hafferl, A. (2013). *Lehrbuch der topographischen Anatomie* (Springer-Verlag).
- [11] Iazzo, P.A. (2010). *Handbook of Cardiac Anatomy, Physiology, and Devices* (Springer Science & Business Media).
- [12] Murphy, P.J. (2005). The fetal circulation. *Contin. Educ. Anaesth. Crit. Care Pain* 5, 107–112.
- [13] Apitz, J. (2013). *Pädiatrische Kardiologie: Erkrankungen des Herzens bei Neugeborenen, Säuglingen, Kindern und Heranwachsenden* (Springer-Verlag).
- [14] Sylva, M., van den Hoff, M.J.B., and Moorman, A.F.M. (2014). Development of the human heart. *Am. J. Med. Genet. A.* 164A, 1347–1371.
- [15] Harvey, R.P. (2002). Patterning the vertebrate heart. *Nat. Rev. Genet.* 3, 544–556.
- [16] Buckingham, M., Meilhac, S., and Zaffran, S. (2005). Building the mammalian heart from two sources of myocardial cells. *Nat. Rev. Genet.* 6, 826–835.
- [17] Moorman, A.F.M., Christoffels, V.M., Anderson, R.H., and van den Hoff, M.J.B. (2007). The heart-forming fields: one or multiple? *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 362, 1257–1265.
- [18] Dyer, L.A., and Kirby, M.L. (2009). The role of secondary heart field in cardiac development. *Dev. Biol.* 336, 137–144.
- [19] Sizarov, A., Ya, J., de Boer, B.A., Lamers, W.H., Christoffels, V.M., and Moorman, A.F.M. (2011). Formation of the building plan of the human heart: morphogenesis, growth, and differentiation. *Circulation* 123, 1125–1135.
- [20] Kelly, R.G., Brown, N.A., and Buckingham, M.E. (2001). The arterial pole of the mouse heart forms from Fgf10-expressing cells in pharyngeal mesoderm. *Dev. Cell* 1, 435–440.
- [21] Vincent, S.D., and Buckingham, M.E. (2010). How to make a heart: the origin and regulation of cardiac progenitor cells. *Curr. Top. Dev. Biol.* 90, 1–41.

- [22] Abu-Issa, R., and Kirby, M.L. (2007). Heart field: from mesoderm to heart tube. *Annu. Rev. Cell Dev. Biol.* 23, 45–68.
- [23] Männer, J. (2000). Cardiac looping in the chick embryo: a morphological review with special reference to terminological and biomechanical aspects of the looping process. *Anat. Rec.* 259, 248–262.
- [24] Bruneau, B.G. (2008). The developmental genetics of congenital heart disease. *Nature* 451, 943–948.
- [25] Christoffels, V.M., Habets, P.E., Franco, D., Campione, M., de Jong, F., Lamers, W.H., Bao, Z.Z., Palmer, S., Biben, C., Harvey, R.P., et al. (2000). Chamber formation and morphogenesis in the developing mammalian heart. *Dev. Biol.* 223, 266–278.
- [26] Moorman, A.F.M., and Christoffels, V.M. (2003). Cardiac chamber formation: development, genes, and evolution. *Physiol. Rev.* 83, 1223–1267.
- [27] Ward, C., Stadt, H., Hutson, M., and Kirby, M.L. (2005). Ablation of the secondary heart field leads to tetralogy of Fallot and pulmonary atresia. *Dev. Biol.* 284, 72–83.
- [28] Zaffran, S., Kelly, R.G., Meilhac, S.M., Buckingham, M.E., and Brown, N.A. (2004). Right ventricular myocardium derives from the anterior heart field. *Circ. Res.* 95, 261–268.
- [29] Verzi, M.P., McCulley, D.J., De Val, S., Dodou, E., and Black, B.L. (2005). The right ventricle, outflow tract, and ventricular septum comprise a restricted expression domain within the secondary/anterior heart field. *Dev. Biol.* 287, 134–145.
- [30] Galli, D., Domínguez, J.N., Zaffran, S., Munk, A., Brown, N.A., and Buckingham, M.E. (2008). Atrial myocardium derives from the posterior region of the second heart field, which acquires left-right identity as *Pitx2c* is expressed. *Dev. Camb. Engl.* 135, 1157–1167.
- [31] Nakajima, Y., Yamagishi, T., Hokari, S., and Nakamura, H. (2000). Mechanisms involved in valvuloseptal endocardial cushion formation in early cardiogenesis: roles of transforming growth factor (TGF)-beta and bone morphogenetic protein (BMP). *Anat. Rec.* 258, 119–127.
- [32] Délot, E.C. (2003). Control of endocardial cushion and cardiac valve maturation by BMP signaling pathways. *Mol. Genet. Metab.* 80, 27–35.
- [33] Hutson, M.R., and Kirby, M.L. (2007). Model systems for the study of heart development and disease. Cardiac neural crest and conotruncal malformations. *Semin. Cell Dev. Biol.* 18, 101–110.
- [34] Ratajska, A., Czarnowska, E., and Cizek, B. (2008). Embryonic development of the proepicardium and coronary vessels. *Int. J. Dev. Biol.* 52, 229–236.
- [35] Kuratani, S.C., and Kirby, M.L. (1992). Migration and distribution of circumpharyngeal crest cells in the chick embryo. Formation of the circumpharyngeal ridge and E/C8+ crest cells in the vertebrate head region. *Anat. Rec.* 234, 263–280.
- [36] Bruneau, B.G. (2010). Epigenetic Regulation of the Cardiovascular System Introduction to a Review Series. *Circ. Res.* 107, 324–326.
- [37] Vallaster, M., Vallaster, C.D., and Wu, S.M. (2012). Epigenetic mechanisms in cardiac development and disease. *Acta Biochim. Biophys. Sin.* 44, 92–102.

- [38] Gilsbach, R., Preissl, S., Grüning, B.A., Schnick, T., Burger, L., Benes, V., Würch, A., Bönisch, U., Günther, S., Backofen, R., et al. (2014). Dynamic DNA methylation orchestrates cardiomyocyte development, maturation and disease. *Nat. Commun.* 5, 5288.
- [39] Chamberlain, A.A., Lin, M., Lister, R.L., Maslov, A.A., Wang, Y., Suzuki, M., Wu, B., Grealley, J.M., Zheng, D., and Zhou, B. (2014). DNA methylation is developmentally regulated for genes essential for cardiogenesis. *J. Am. Heart Assoc.* 3, e000976.
- [40] Nerem, R.M., Harrison, D.G., Taylor, W.R., and Alexander, R.W. (1993). Hemodynamics and vascular endothelial biology. *J. Cardiovasc. Pharmacol.* 21 Suppl 1, S6–S10.
- [41] Takahashi, M., Ishida, T., Traub, O., Corson, M.A., and Berk, B.C. (1997). Mechanotransduction in endothelial cells: temporal signaling events in response to shear stress. *J. Vasc. Res.* 34, 212–219.
- [42] Taber, L.A. (1998). Mechanical aspects of cardiac development. *Prog. Biophys. Mol. Biol.* 69, 237–255.
- [43] Epstein, J.A. (2010). Franklin H. Epstein Lecture. Cardiac development and implications for heart disease. *N. Engl. J. Med.* 363, 1638–1647.
- [44] Hoffman, J.I.E., Kaplan, S., and Liberthson, R.R. (2004). Prevalence of congenital heart disease. *Am. Heart J.* 147, 425–439.
- [45] Rubart, M., and Zipes, D.P. (2005). Mechanisms of sudden cardiac death. *J. Clin. Invest.* 115, 2305–2315.
- [46] Mitchell, S.C., Korones, S.B., and Berendes, H.W. (1971). Congenital heart disease in 56,109 births. Incidence and natural history. *Circulation* 43, 323–332.
- [47] Fahed, A.C., Gelb, B.D., Seidman, J.G., and Seidman, C.E. (2013). Genetics of congenital heart disease: the glass half empty. *Circ. Res.* 112, 707–720.
- [48] Talner, N.S. (1998). Report of the New England Regional Infant Cardiac Program, by Donald C. Fyler, MD, *Pediatrics*, 1980;65(suppl):375–461. *Pediatrics* 102, 258–259.
- [49] Oster, M.E., Lee, K.A., Honein, M.A., Riehle-Colarusso, T., Shin, M., and Correa, A. (2013). Temporal trends in survival among infants with critical congenital heart defects. *Pediatrics* 131, e1502–e1508.
- [50] Yang, Q., Chen, H., Correa, A., Devine, O., Mathews, T. j., and Honein, M.A. (2006). Racial differences in infant mortality attributable to birth defects in the United States, 1989–2002. *Birt. Defects Res. A. Clin. Mol. Teratol.* 76, 706–713.
- [51] Peter W G Tennant, M.S.P. (2010). 20-Year survival of children born with congenital anomalies: A population-based study. *Lancet* 375, 649–656.
- [52] Roger, V.L., Go, A.S., Lloyd-Jones, D.M., Benjamin, E.J., Berry, J.D., Borden, W.B., Bravata, D.M., Dai, S., Ford, E.S., Fox, C.S., et al. (2012). Heart Disease and Stroke Statistics-2012 Update A Report From the American Heart Association. *Circulation* 125, E2–E220.
- [53] Coding Committee of the Association for European Paediatric Cardiology (2002). The European Paediatric Cardiac Code: the first revision. *Cardiol. Young* 12 Suppl 2, 1–211.

- [54] Borth-Bruhns, T., and Eichler, A. (2013). *Pädiatrische Kardiologie* (Springer-Verlag).
- [55] Down, J.L. (1995). Observations on an ethnic classification of idiots. 1866. *Ment. Retard.* 33, 54–56.
- [56] J, L., M, G., and R, T. (1959). Etude des chromosomes somatiques de neuf enfants mongoliens. *Comptes Rendus Hebd. Séances Académie Sci.* 248, 1721–1722.
- [57] Brekke, V.G. (1953). Congenital aortic atresia and hypoplasia of the aortic orifice; case reports on two members of a family. *Am. Heart J.* 45, 925–930.
- [58] Papagiannis, J., Van Praagh, R., Schwint, O., D’Orsogna, L., Qureshi, F., Reynolds, J., Kallfelz, C., and Nozar, J. (2001). Congenital left ventricular aneurysm: clinical, imaging, pathologic, and surgical findings in seven new cases. *Am. Heart J.* 141, 491–499.
- [59] Jenkins, K.J., Correa, A., Feinstein, J.A., Botto, L., Britt, A.E., Daniels, S.R., Elixson, M., Warnes, C.A., Webb, C.L., and American Heart Association Council on Cardiovascular Disease in the Young (2007). Noninherited risk factors and congenital cardiovascular defects: current knowledge: a scientific statement from the American Heart Association Council on Cardiovascular Disease in the Young: endorsed by the American Academy of Pediatrics. *Circulation* 115, 2995–3014.
- [60] Marelli, A.J., Mackie, A.S., Ionescu-Iltu, R., Rahme, E., and Pilote, L. (2007). Congenital heart disease in the general population: changing prevalence and age distribution. *Circulation* 115, 163–172.
- [61] Billett, J., Majeed, A., Gatzoulis, M., and Cowie, M. (2008). Trends in hospital admissions, in-hospital case fatality and population mortality from congenital heart disease in England, 1994 to 2004. *Heart Br. Card. Soc.* 94, 342–348.
- [62] Oyen, N., Poulsen, G., Boyd, H.A., Wohlfahrt, J., Jensen, P.K.A., and Melbye, M. (2009). National time trends in congenital heart defects, Denmark, 1977-2005. *Am. Heart J.* 157, 467–473.e1.
- [63] Hoffman, J.I.E., and Kaplan, S. (2002). The incidence of congenital heart disease. *J. Am. Coll. Cardiol.* 39, 1890–1900.
- [64] Bernier, P.-L., Stefanescu, A., Samoukovic, G., and Tchervenkov, C.I. (2010). The challenge of congenital heart disease worldwide: epidemiologic and demographic facts. *Semin. Thorac. Cardiovasc. Surg. Pediatr. Card. Surg. Annu.* 13, 26–34.
- [65] Hoffman, J.I.E. (1995). Incidence of congenital heart disease: I. Postnatal incidence. *Pediatr. Cardiol.* 16, 103–113.
- [66] Øyen, N., Poulsen, G., Wohlfahrt, J., Boyd, H.A., Jensen, P.K.A., and Melbye, M. (2010). Recurrence of Discordant Congenital Heart Defects in Families. *Circ. Cardiovasc. Genet.* 3, 122–128.

- [67] Fesslova, V., Brankovic, J., Lalatta, F., Villa, L., Meli, V., Piazza, L., Ricci, C., Fesslova, V., Brankovic, J., Lalatta, F., et al. (2011). Recurrence of Congenital Heart Disease in Cases with Familial Risk Screened Prenatally by Echocardiography, Recurrence of Congenital Heart Disease in Cases with Familial Risk Screened Prenatally by Echocardiography. *J. Pregnancy J. Pregnancy* 2011, 2011, e368067.
- [68] van der Bom, T., Zomer, A.C., Zwinderman, A.H., Meijboom, F.J., Bouma, B.J., and Mulder, B.J.M. (2011). The changing epidemiology of congenital heart disease. *Nat. Rev. Cardiol.* 8, 50–60.
- [69] Blue, G.M., Kirk, E.P., Sholler, G.F., Harvey, R.P., and Winlaw, D.S. (2012). Congenital heart disease: current knowledge about causes and inheritance. *Med. J. Aust.* 197.
- [70] Hardin, J., Carmichael, S.L., Selvin, S., Lammer, E.J., and Shaw, G.M. (2009). Increased prevalence of cardiovascular defects among 56,709 California twin pairs. *Am. J. Med. Genet. A.* 149A, 877–886.
- [71] Lewin, M.B., McBride, K.L., Pignatelli, R., Fernbach, S., Combes, A., Menesses, A., Lam, W., Bezold, L.I., Kaplan, N., Towbin, J.A., et al. (2004). Echocardiographic Evaluation of Asymptomatic Parental and Sibling Cardiovascular Anomalies Associated With Congenital Left Ventricular Outflow Tract Lesions. *Pediatrics* 114, 691–696.
- [72] Gill, H.K., Splitt, M., Sharland, G.K., and Simpson, J.M. (2003). Patterns of recurrence of congenital heart disease: an analysis of 6,640 consecutive pregnancies evaluated by detailed fetal echocardiography. *J. Am. Coll. Cardiol.* 42, 923–929.
- [73] Wu Ct, null, and Morris, J.R. (2001). Genes, genetics, and epigenetics: a correspondence. *Science* 293, 1103–1105.
- [74] Zaidi, S., Choi, M., Wakimoto, H., Ma, L., Jiang, J., Overton, J.D., Romano-Adesman, A., Bjornson, R.D., Breitbart, R.E., Brown, K.K., et al. (2013). De novo mutations in histone-modifying genes in congenital heart disease. *Nature* 498, 220–223.
- [75] Nora, J.J. (1968). Multifactorial inheritance hypothesis for the etiology of congenital heart diseases. The genetic-environmental interaction. *Circulation* 38, 604–617.
- [76] Nora, J.J., McGill, C.W., and McNamara, D.G. (1970). Empiric recurrence risks in common and uncommon congenital heart lesions. *Teratology* 3, 325–330.
- [77] Pierpont, M.E., Basson, C.T., Benson, D.W., Gelb, B.D., Giglia, T.M., Goldmuntz, E., McGee, G., Sable, C.A., Srivastava, D., Webb, C.L., et al. (2007). Genetic basis for congenital heart defects: current knowledge: a scientific statement from the American Heart Association Congenital Cardiac Defects Committee, Council on Cardiovascular Disease in the Young: endorsed by the American Academy of Pediatrics. *Circulation* 115, 3015–3038.
- [78] Wessels, M.W., and Willems, P.J. (2010). Genetic factors in non-syndromic congenital heart malformations. *Clin. Genet.* 78, 103–123.
- [79] Lalani, S.R., and Belmont, J.W. (2014). Genetic Basis of Congenital Cardiovascular Malformations. *Eur. J. Med. Genet.* 57, 402–413.

- [80] Roos-Hesselink, J.W., Kerstjens-Frederikse, W.S., Meijboom, F.J., and Pieper, P.G. (2005). Inheritance of congenital heart disease. *Neth. Heart J.* 13, 88–91.
- [81] Hartman, R.J., Rasmussen, S.A., Botto, L.D., Riehle-Colarusso, T., Martin, C.L., Cragan, J.D., Shin, M., and Correa, A. (2011). The contribution of chromosomal abnormalities to congenital heart defects: a population-based study. *Pediatr. Cardiol.* 32, 1147–1157.
- [82] Antonarakis, S.E., Lyle, R., Dermitzakis, E.T., Reymond, A., and Deutsch, S. (2004). Chromosome 21 and down syndrome: from genomics to pathophysiology. *Nat. Rev. Genet.* 5, 725–738.
- [83] Bondy, C.A. (2009). Turner syndrome 2008. *Horm. Res.* 71 Suppl 1, 52–56.
- [84] Pont, S.J., Robbins, J.M., Bird, T.M., Gibson, J.B., Cleves, M.A., Tilford, J.M., and Aitken, M.E. (2006). Congenital malformations among liveborn infants with trisomies 18 and 13. *Am. J. Med. Genet. A.* 140, 1749–1756.
- [85] Lahm, H., Schön, P., Doppler, S., Dreßen, M., Cleuziou, J., Deutsch, M.-A., Ewert, P., Lange, R., and Krane, M. (2015). Tetralogy of Fallot and Hypoplastic Left Heart Syndrome - Complex Clinical Phenotypes Meet Complex Genetic Networks. *Curr. Genomics* 16, 141–158.
- [86] Goldmuntz, E. (2005). DiGeorge syndrome: new insights. *Clin. Perinatol.* 32, 963–978, ix – x.
- [87] Greenway, S.C., Pereira, A.C., Lin, J.C., DePalma, S.R., Israel, S.J., Mesquita, S.M., Ergul, E., Conta, J.H., Korn, J.M., McCarroll, S.A., et al. (2009). De novo copy number variants identify new genes and loci in isolated sporadic tetralogy of Fallot. *Nat. Genet.* 41, 931–935.
- [88] Silversides, C.K., Lionel, A.C., Costain, G., Merico, D., Migita, O., Liu, B., Yuen, T., Rickaby, J., Thiruvahindrapuram, B., Marshall, C.R., et al. (2012). Rare copy number variations in adults with tetralogy of Fallot implicate novel risk gene pathways. *PLoS Genet.* 8, e1002843.
- [89] Hitz, M.-P., Lemieux-Perreault, L.-P., Marshall, C., Feroz-Zada, Y., Davies, R., Yang, S.W., Lionel, A.C., D'Amours, G., Lemyre, E., Cullum, R., et al. (2012). Rare copy number variants contribute to congenital left-sided heart disease. *PLoS Genet.* 8, e1002903.
- [90] Payne, A.R., Chang, S.-W., Koenig, S.N., Zinn, A.R., and Garg, V. (2012). Submicroscopic chromosomal copy number variations identified in children with hypoplastic left heart syndrome. *Pediatr. Cardiol.* 33, 757–763.
- [91] Erdogan, F., Larsen, L.A., Zhang, L., Tümer, Z., Tommerup, N., Chen, W., Jacobsen, J.R., Schubert, M., Jurkatis, J., Tzschach, A., et al. (2008). High frequency of submicroscopic genomic aberrations detected by tiling path array comparative genome hybridisation in patients with isolated congenital heart disease. *J. Med. Genet.* 45, 704–709.
- [92] Priest, J.R., Girirajan, S., Vu, T.H., Olson, A., Eichler, E.E., and Portman, M.A. (2012). Rare copy number variants in isolated sporadic and syndromic atrioventricular septal defects. *Am. J. Med. Genet. A.* 158A, 1279–1284.

- [93] Soemedi, R., Wilson, I.J., Bentham, J., Darlay, R., Töpf, A., Zelenika, D., Cosgrove, C., Setchfield, K., Thornborough, C., Granados-Riveron, J., et al. (2012). Contribution of Global Rare Copy-Number Variants to the Risk of Sporadic Congenital Heart Disease. *Am. J. Hum. Genet.* 91, 489–501.
- [94] Tartaglia, M., Mehler, E.L., Goldberg, R., Zampino, G., Brunner, H.G., Kremer, H., van der Burgt, I., Crosby, A.H., Ion, A., Jeffery, S., et al. (2001). Mutations in PTPN11, encoding the protein tyrosine phosphatase SHP-2, cause Noonan syndrome. *Nat. Genet.* 29, 465–468.
- [95] Roberts, A.E., Araki, T., Swanson, K.D., Montgomery, K.T., Schiripo, T.A., Joshi, V.A., Li, L., Yassin, Y., Tamburino, A.M., Neel, B.G., et al. (2007). Germline gain-of-function mutations in SOS1 cause Noonan syndrome. *Nat. Genet.* 39, 70–74.
- [96] Razzaque, M.A., Nishizawa, T., Komoike, Y., Yagi, H., Furutani, M., Amo, R., Kamisago, M., Momma, K., Katayama, H., Nakagawa, M., et al. (2007). Germline gain-of-function mutations in RAF1 cause Noonan syndrome. *Nat. Genet.* 39, 1013–1017.
- [97] Marino, B., Digilio, M.C., Toscano, A., Giannotti, A., and Dallapiccola, B. (1999). Congenital heart diseases in children with Noonan syndrome: An expanded cardiac spectrum with high prevalence of atrioventricular canal. *J. Pediatr.* 135, 703–706.
- [98] Digilio, M., and Marino, B. (2001). Clinical manifestations of Noonan syndrome. *Images Paediatr. Cardiol.* 3, 19–30.
- [99] Oda, T., Elkahloun, A.G., Pike, B.L., Okajima, K., Krantz, I.D., Genin, A., Piccoli, D.A., Meltzer, P.S., Spinner, N.B., Collins, F.S., et al. (1997). Mutations in the human Jagged1 gene are responsible for Alagille syndrome. *Nat. Genet.* 16, 235–242.
- [100] Li, L., Krantz, I.D., Deng, Y., Genin, A., Banta, A.B., Collins, C.C., Qi, M., Trask, B.J., Kuo, W.L., Cochran, J., et al. (1997). Alagille syndrome is caused by mutations in human Jagged1, which encodes a ligand for Notch1. *Nat. Genet.* 16, 243–251.
- [101] Subramaniam, P., Knisely, A., Portmann, B., Qureshi, S.A., Aclimandos, W.A., Karani, J.B., and Baker, A.J. (2011). Diagnosis of Alagille syndrome-25 years of experience at King's College Hospital. *J. Pediatr. Gastroenterol. Nutr.* 52, 84–89.
- [102] Warthen, D.M., Moore, E.C., Kamath, B.M., Morrissette, J.J.D., Sanchez-Lara, P.A., Sanchez, P., Piccoli, D.A., Krantz, I.D., and Spinner, N.B. (2006). Jagged1 (JAG1) mutations in Alagille syndrome: increasing the mutation detection rate. *Hum. Mutat.* 27, 436–443.
- [103] Manzke, H., Lehmann, K., Klopocki, E., and Caliebe, A. (2008). Catel-Manzke syndrome: two new patients and a critical review of the literature. *Eur. J. Med. Genet.* 51, 452–465.
- [104] Ehmke, N., Caliebe, A., Koenig, R., Kant, S.G., Stark, Z., Cormier-Daire, V., Wieczorek, D., Gillissen-Kaesbach, G., Hoff, K., Kawalia, A., et al. (2014). Homozygous and Compound-Heterozygous Mutations in TGDS Cause Catel-Manzke Syndrome. *Am. J. Hum. Genet.* 95, 763–770.

- [105] Schott, J.J., Benson, D.W., Basson, C.T., Pease, W., Silberbach, G.M., Moak, J.P., Maron, B.J., Seidman, C.E., and Seidman, J.G. (1998). Congenital heart disease caused by mutations in the transcription factor NKX2-5. *Science* 281, 108–111.
- [106] Elliott, D.A., Kirk, E.P., Yeoh, T., Chandar, S., McKenzie, F., Taylor, P., Grossfeld, P., Fatkin, D., Jones, O., Hayes, P., et al. (2003). Cardiac homeobox gene NKX2-5 mutations and congenital heart disease: associations with atrial septal defect and hypoplastic left heart syndrome. *J. Am. Coll. Cardiol.* 41, 2072–2076.
- [107] McElhinney, D.B., Geiger, E., Blinder, J., Benson, D.W., and Goldmuntz, E. (2003). NKX2.5 mutations in patients with congenital heart disease. *J. Am. Coll. Cardiol.* 42, 1650–1655.
- [108] Sarkozy, A., Conti, E., Neri, C., D’Agostino, R., Digilio, M.C., Esposito, G., Toscano, A., Marino, B., Pizzuti, A., and Dallapiccola, B. (2005). Spectrum of atrial septal defects associated with mutations of NKX2.5 and GATA4 transcription factors. *J. Med. Genet.* 42, e16.
- [109] Tomita-Mitchell, A., Maslen, C.L., Morris, C.D., Garg, V., and Goldmuntz, E. (2007). GATA4 sequence variants in patients with congenital heart disease. *J. Med. Genet.* 44, 779–783.
- [110] Garg, V., Kathiriyia, I.S., Barnes, R., Schluterman, M.K., King, I.N., Butler, C.A., Rothrock, C.R., Eapen, R.S., Hirayama-Yamada, K., Joo, K., et al. (2003). GATA4 mutations cause human congenital heart defects and reveal an interaction with TBX5. *Nature* 424, 443–447.
- [111] Peng, T., Wang, L., Zhou, S.-F., and Li, X. (2010). Mutations of the GATA4 and NKX2.5 genes in Chinese pediatric patients with non-familial congenital heart disease. *Genetica* 138, 1231–1240.
- [112] Biben, C., Weber, R., Kesteven, S., Stanley, E., McDonald, L., Elliott, D.A., Barnett, L., Köentgen, F., Robb, L., Feneley, M., et al. (2000). Cardiac septal and valvular dysmorphogenesis in mice heterozygous for mutations in the homeobox gene *Nkx2-5*. *Circ. Res.* 87, 888–895.
- [113] Bruneau, B.G., Nemer, G., Schmitt, J.P., Charron, F., Robitaille, L., Caron, S., Conner, D.A., Gessler, M., Nemer, M., Seidman, C.E., et al. (2001a). A murine model of Holt-Oram syndrome defines roles of the T-box transcription factor *Tbx5* in cardiogenesis and disease. *Cell* 106, 709–721.
- [114] Laforest, B., and Nemer, M. (2011). GATA5 interacts with GATA4 and GATA6 in outflow tract development. *Dev. Biol.* 358, 368–378.
- [115] Huang, R.-T., Xue, S., Xu, Y.-J., and Yang, Y.-Q. (2013). Somatic mutations in the GATA6 gene underlie sporadic tetralogy of Fallot. *Int. J. Mol. Med.* 31, 51–58.
- [116] Gollob, M.H., Jones, D.L., Krahn, A.D., Danis, L., Gong, X.-Q., Shao, Q., Liu, X., Veinot, J.P., Tang, A.S.L., Stewart, A.F.R., et al. (2006). Somatic Mutations in the Connexin 40 Gene (*GJA5*) in Atrial Fibrillation. *N. Engl. J. Med.* 354, 2677–2688.
- [117] Thibodeau, I.L., Xu, J., Li, Q., Liu, G., Lam, K., Veinot, J.P., Birnie, D.H., Jones, D.L., Krahn, A.D., Lemery, R., et al. (2010). Paradigm of genetic mosaicism and lone atrial fibrillation: physiological characterization of a connexin 43-deletion mutant identified from atrial tissue. *Circulation* 122, 236–244.

- [118] Choi, M., Scholl, U.I., Ji, W., Liu, T., Tikhonova, I.R., Zumbo, P., Nayir, A., Bakkaloğlu, A., Özen, S., Sanjad, S., et al. (2009a). Genetic diagnosis by whole exome capture and massively parallel DNA sequencing. *Proc. Natl. Acad. Sci.* 106, 19096–19101.
- [119] Sanders, S.J., Murtha, M.T., Gupta, A.R., Murdoch, J.D., Raubeson, M.J., Willsey, A.J., Ercan-Sencicek, A.G., DiLullo, N.M., Parikshak, N.N., Stein, J.L., et al. (2012). De novo mutations revealed by whole-exome sequencing are strongly associated with autism. *Nature* 485, 237–241.
- [120] Waddington CH. The epigenotype. *Endeavour.* 1942;1:18–20.
- [121] Kim, J.K., Samaranyake, M., and Pradhan, S. (2009). Epigenetic mechanisms in mammals. *Cell. Mol. Life Sci.* 66, 596–612.
- [122] Heyn, H., and Esteller, M. (2012). DNA methylation profiling in the clinic: applications and challenges. *Nat. Rev. Genet.* 13, 679–692.
- [123] Reik, W. (2007). Stability and flexibility of epigenetic gene regulation in mammalian development. *Nature* 447, 425–432.
- [124] Law, J.A., and Jacobsen, S.E. (2010). Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nat. Rev. Genet.* 11, 204–220.
- [125] Kota, S.K., and Feil, R. (2010). Epigenetic transitions in germ cell development and meiosis. *Dev. Cell* 19, 675–686.
- [126] Feil, R., and Fraga, M.F. (2011). Epigenetics and the environment: emerging patterns and implications. *Nat. Rev. Genet.* 13, 97–109.
- [127] Marsit, C.J. (2015). Influence of environmental exposure on human epigenetic regulation. *J. Exp. Biol.* 218, 71–79.
- [128] Srivastava, D. (2001). Genetic assembly of the heart: implications for congenital heart disease. *Annu. Rev. Physiol.* 63, 451–469.
- [129] Correa, A., Gilboa, S.M., Besser, L.M., Botto, L.D., Moore, C.A., Hobbs, C.A., Cleves, M.A., Riehle-Colarusso, T.J., Waller, D.K., and Reece, E.A. (2008). Diabetes mellitus and birth defects. *Am. J. Obstet. Gynecol.* 199, 237.e1–e9.
- [130] Bánhidly, F., Acs, N., Puhó, E.H., and Czeizel, A.E. (2010). Congenital abnormalities in the offspring of pregnant women with type 1, type 2 and gestational diabetes mellitus: a population-based case-control study. *Congenit. Anom.* 50, 115–121.
- [131] Correa, A., Gilboa, S.M., Botto, L.D., Moore, C.A., Hobbs, C.A., Cleves, M.A., Riehle-Colarusso, T.J., Waller, D.K., Reece, E.A., and National Birth Defects Prevention Study (2012). Lack of periconceptional vitamins or supplements that contain folic acid and diabetes mellitus-associated birth defects. *Am. J. Obstet. Gynecol.* 206, 218.e1–e13.
- [132] Ling, C., Del Guerra, S., Lupi, R., Rönn, T., Granhall, C., Luthman, H., Masiello, P., Marchetti, P., Groop, L., and Del Prato, S. (2008). Epigenetic regulation of PPARGC1A in human type 2 diabetic islets and effect on insulin secretion. *Diabetologia* 51, 615–622.
- [133] Zhao, J., Goldberg, J., Bremner, J.D., and Vaccarino, V. (2012). Global DNA Methylation Is Associated With Insulin Resistance A Monozygotic Twin Study. *Diabetes* 61, 542–546.

- [134] Toperoff, G., Aran, D., Kark, J.D., Rosenberg, M., Dubnikov, T., Nissan, B., Wainstein, J., Friedlander, Y., Levy-Lahad, E., Glaser, B., et al. (2012). Genome-wide survey reveals predisposing diabetes type 2-related DNA methylation variations in human peripheral blood. *Hum. Mol. Genet.* 21, 371–383.
- [135] Kopf, P.G., and Walker, M.K. (2009). Overview of developmental heart defects by dioxins, PCBs, and pesticides. *J. Environ. Sci. Health Part C Environ. Carcinog. Ecotoxicol. Rev.* 27, 276–285.
- [136] Ferencz, C., Boughman, J.A., Neill, C.A., Brenner, J.I., and Perry, L.W. (1989). Congenital cardiovascular malformations: questions on inheritance. Baltimore-Washington Infant Study Group. *J. Am. Coll. Cardiol.* 14, 756–763.
- [137] Lipshultz, S.E., Frassica, J.J., and Orav, E.J. (1991). Cardiovascular abnormalities in infants prenatally exposed to cocaine. *J. Pediatr.* 118, 44–51.
- [138] Alverson, C.J., Strickland, M.J., Gilboa, S.M., and Correa, A. (2011). Maternal smoking and congenital heart defects in the Baltimore-Washington Infant Study. *Pediatrics* 127, e647–e653.
- [139] Mateja, W.A., Nelson, D.B., Kroelinger, C.D., Ruzek, S., and Segal, J. (2012). The association between maternal alcohol use and smoking in early pregnancy and congenital cardiac defects. *J. Womens Health* 2002 21, 26–34.
- [140] Liu, Y., Balaraman, Y., Wang, G., Nephew, K.P., and Zhou, F.C. (2009). Alcohol exposure alters DNA methylation profiles in mouse embryos at early neurulation. *Epigenetics* 4, 500–511.
- [141] Kazaura, M., Lie, R.T., and Skjaerven, R. (2004). Paternal age and the risk of birth defects in Norway. *Ann. Epidemiol.* 14, 566–570.
- [142] Materna-Kiryluk, A., Wiśniewska, K., Badura-Stronka, M., Mejnartowicz, J., Wieckowska, B., Balcar-Boroń, A., Czerwionka-Szaflarska, M., Gajewska, E., Godula-Stuglik, U., Krawczyński, M., et al. (2009). Parental age as a risk factor for isolated congenital malformations in a Polish population. *Paediatr. Perinat. Epidemiol.* 23, 29–40.
- [143] Miller, A., Riehle-Colarusso, T., Siffel, C., Frías, J.L., and Correa, A. (2011). Maternal age and prevalence of isolated congenital heart defects in an urban area of the United States. *Am. J. Med. Genet. A.* 155A, 2137–2145.
- [144] Schulkey, C.E., Regmi, S.D., Magnan, R.A., Danzo, M.T., Luther, H., Hutchinson, A.K., Panzer, A.A., Grady, M.M., Wilson, D.B., and Jay, P.Y. (2015). The maternal-age-associated risk of congenital heart disease is modifiable. *Nature* 520, 230–233.
- [145] Fraga, M.F., Ballestar, E., Paz, M.F., Ropero, S., Setien, F., Ballestar, M.L., Heine-Suñer, D., Cigudosa, J.C., Urioste, M., Benitez, J., et al. (2005). Epigenetic differences arise during the lifetime of monozygotic twins. *Proc. Natl. Acad. Sci. U. S. A.* 102, 10604–10609.
- [146] Kwabi-Addo, B., Chung, W., Shen, L., Ittmann, M., Wheeler, T., Jelinek, J., and Issa, J.-P.J. (2007). Age-Related DNA Methylation Changes in Normal Human Prostate Tissues. *Clin. Cancer Res.* 13, 3796–3802.

- [147] Bjornsson, H.T., Sigurdsson, M.I., Fallin, M.D., Irizarry, R.A., Aspelund, T., Cui, H., Yu, W., Rongione, M.A., Ekström, T.J., Harris, T.B., et al. (2008). Intra-individual change over time in DNA methylation with familial clustering. *JAMA* 299, 2877–2883.
- [148] Boks, M.P., Derks, E.M., Weisenberger, D.J., Strengman, E., Janson, E., Sommer, I.E., Kahn, R.S., and Ophoff, R.A. (2009). The relationship of DNA methylation with age, gender and genotype in twins and healthy controls. *PLoS One* 4, e6767.
- [149] Christensen, B.C., Houseman, E.A., Marsit, C.J., Zheng, S., Wrensch, M.R., Wiemels, J.L., Nelson, H.H., Karagas, M.R., Padbury, J.F., Bueno, R., et al. (2009). Aging and environmental exposures alter tissue-specific DNA methylation dependent upon CpG island context. *PLoS Genet.* 5, e1000602.
- [150] Rakyan, V.K., Down, T.A., Maslau, S., Andrew, T., Yang, T.-P., Beyan, H., Whittaker, P., McCann, O.T., Finer, S., Valdes, A.M., et al. (2010). Human aging-associated DNA hypermethylation occurs preferentially at bivalent chromatin domains. *Genome Res.* 20, 434–439.
- [151] Talens, R.P., Christensen, K., Putter, H., Willemsen, G., Christiansen, L., Kremer, D., Suchiman, H.E.D., Slagboom, P.E., Boomsma, D.I., and Heijmans, B.T. (2012). Epigenetic variation during the adult lifespan: cross-sectional and longitudinal data on monozygotic twin pairs. *Aging Cell* 11, 694–703.
- [152] Noreen, F., Rösli, M., Gaj, P., Pietrzak, J., Weis, S., Urfer, P., Regula, J., Schär, P., and Truninger, K. (2014). Modulation of Age- and Cancer-Associated DNA Methylation Change in the Healthy Colon by Aspirin and Lifestyle. *J. Natl. Cancer Inst.* 106, dju161.
- [153] Jeltsch, A. (2002). Beyond Watson and Crick: DNA methylation and molecular enzymology of DNA methyltransferases. *ChemBiochem Eur. J. Chem. Biol.* 3, 274–293.
- [154] Lister, R., and Ecker, J.R. (2009). Finding the fifth base: genome-wide sequencing of cytosine methylation. *Genome Res.* 19, 959–966.
- [155] Jabbari, K., and Bernardi, G. (2004). Cytosine methylation and CpG, TpG (CpA) and TpA frequencies. *Gene* 333, 143–149.
- [156] Ramsahoye, B.H., Biniszkiwicz, D., Lyko, F., Clark, V., Bird, A.P., and Jaenisch, R. (2000). Non-CpG methylation is prevalent in embryonic stem cells and may be mediated by DNA methyltransferase 3a. *Proc. Natl. Acad. Sci. U. S. A.* 97, 5237–5242.
- [157] Bestor, T.H., and Verdine, G.L. (1994). DNA methyltransferases. *Curr. Opin. Cell Biol.* 6, 380–389.
- [158] Vilkaitis, G., Suetake, I., Klimasauskas, S., and Tajima, S. (2005). Processive methylation of hemimethylated CpG sites by mouse Dnmt1 DNA methyltransferase. *J. Biol. Chem.* 280, 64–72.
- [159] Okano, M., Bell, D.W., Haber, D.A., and Li, E. (1999). DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* 99, 247–257.
- [160] Wu, S.C., and Zhang, Y. (2010). Active DNA demethylation: many roads lead to Rome. *Nat. Rev. Mol. Cell Biol.* 11, 607–620.

- [161] Takai, D., and Jones, P.A. (2002). Comprehensive analysis of CpG islands in human chromosomes 21 and 22. *Proc. Natl. Acad. Sci.* 99, 3740–3745.
- [162] Larsen, F., Gundersen, G., Lopez, R., and Prydz, H. (1992). CpG islands as gene markers in the human genome. *Genomics* 13, 1095–1107.
- [163] Zhu, J., He, F., Hu, S., and Yu, J. (2008). On the nature of human housekeeping genes. *Trends Genet.* TIG 24, 481–484.
- [164] Holliday, R., and Pugh, J.E. (1975). DNA modification mechanisms and gene activity during development. *Science* 187, 226–232.
- [165] Wolffe, A.P., and Matzke, M.A. (1999). Epigenetics: regulation through repression. *Science* 286, 481–486.
- [166] Jones, P.A., and Takai, D. (2001). The role of DNA methylation in mammalian epigenetics. *Science* 293, 1068–1070.
- [167] Wade, P.A. (2001). Methyl CpG-binding proteins and transcriptional repression. *BioEssays News Rev. Mol. Cell. Dev. Biol.* 23, 1131–1137.
- [168] Tate, P.H., and Bird, A.P. (1993). Effects of DNA methylation on DNA-binding proteins and gene expression. *Curr. Opin. Genet. Dev.* 3, 226–231.
- [169] Kim, J., Kollhoff, A., Bergmann, A., and Stubbs, L. (2003). Methylation-sensitive binding of transcription factor YY1 to an insulator sequence within the paternally expressed imprinted gene, *Peg3*. *Hum. Mol. Genet.* 12, 233–245.
- [170] Perini, G., Diolaiti, D., Porro, A., and Della Valle, G. (2005). In vivo transcriptional regulation of N-Myc target genes is controlled by E-box methylation. *Proc. Natl. Acad. Sci. U. S. A.* 102, 12117–12122.
- [171] Choy, M.-K., Movassagh, M., Goh, H.-G., Bennett, M.R., Down, T.A., and Foo, R.S. (2010). Genome-wide conserved consensus transcription factor binding motifs are hyper-methylated. *BMC Genomics* 11, 519.
- [172] Rakan, V.K., Down, T.A., Balding, D.J., and Beck, S. (2011). Epigenome-wide association studies for common human diseases. *Nat. Rev. Genet.* 12, 529–541.
- [173] Doi, A., Park, I.-H., Wen, B., Murakami, P., Aryee, M.J., Irizarry, R., Herb, B., Ladd-Acosta, C., Rho, J., Loewer, S., et al. (2009). Differential methylation of tissue- and cancer-specific CpG island shores distinguishes human induced pluripotent stem cells, embryonic stem cells and fibroblasts. *Nat. Genet.* 41, 1350–1353.
- [174] Bird, A. (2002). DNA methylation patterns and epigenetic memory. *Genes Dev.* 16, 6–21.
- [175] Pennacchio, L.A., Bickmore, W., Dean, A., Nobrega, M.A., and Bejerano, G. (2013). Enhancers: five essential questions. *Nat. Rev. Genet.* 14, 288–295.
- [176] Blattler, A., Yao, L., Witt, H., Guo, Y., Nicolet, C.M., Berman, B.P., and Farnham, P.J. (2014). Global loss of DNA methylation uncovers intronic enhancers in genes showing expression changes. *Genome Biol.* 15, 469.

- [177] Movassagh, M., Choy, M.-K., Knowles, D.A., Cordeddu, L., Haider, S., Down, T., Siggins, L., Vujic, A., Simeoni, I., Penkett, C., et al. (2011). Distinct epigenomic features in end-stage failing human hearts. *Circulation* 124, 2411–2422.
- [178] Haas, J., Frese, K.S., Park, Y.J., Keller, A., Vogel, B., Lindroth, A.M., Weichenhan, D., Franke, J., Fischer, S., Bauer, A., et al. (2013). Alterations in cardiac DNA methylation in human dilated cardiomyopathy. *EMBO Mol. Med.* 5, 413–429.
- [179] Sheng, W., Qian, Y., Wang, H., Ma, X., Zhang, P., Diao, L., An, Q., Chen, L., Ma, D., and Huang, G. (2013). DNA methylation status of NKX2-5, GATA4 and HAND1 in patients with tetralogy of fallot. *BMC Med. Genomics* 6, 46.
- [180] Obermann-Borst, S.A., van Driel, L.M.J.W., Helbing, W.A., de Jonge, R., Wildhagen, M.F., Steegers, E.A.P., and Steegers-Theunissen, R.P.M. (2011). Congenital heart defects and biomarkers of methylation in children: a case-control study. *Eur. J. Clin. Invest.* 41, 143–150.
- [181] Tchervenkov, C.I., Jacobs, M.L., and Tahta, S.A. (2000). Congenital Heart Surgery Nomenclature and Database Project: hypoplastic left heart syndrome. *Ann. Thorac. Surg.* 69, S170–S179.
- [182] Loffredo, C.A. (2000). Epidemiology of cardiovascular malformations: prevalence and risk factors. *Am. J. Med. Genet.* 97, 319–325.
- [183] Gordon, B.M., Rodriguez, S., Lee, M., and Chang, R.-K. (2008). Decreasing number of deaths of infants with hypoplastic left heart syndrome. *J. Pediatr.* 153, 354–358.
- [184] Karamlou, T., Diggs, B.S., Ungerleider, R.M., and Welke, K.F. (2010). Evolution of treatment options and outcomes for hypoplastic left heart syndrome over an 18-year period. *J. Thorac. Cardiovasc. Surg.* 139, 119–126; discussion 126–127.
- [185] Fixler, D.E., Nembhard, W.N., Salemi, J.L., Ethen, M.K., and Canfield, M.A. (2010). Mortality in first 5 years in infants with functional single ventricle born in Texas, 1996 to 2003. *Circulation* 121, 644–650.
- [186] Reller, M.D., Strickland, M.J., Riehle-Colarusso, T., Mahle, W.T., and Correa, A. (2008). Prevalence of congenital heart defects in metropolitan Atlanta, 1998-2005. *J. Pediatr.* 153, 807–813.
- [187] Lindinger, A., Schwedler, G., and Hense, H.-W. (2010). Prevalence of congenital heart defects in newborns in Germany: Results of the first registration year of the PAN Study (July 2006 to June 2007). *Klin. Pädiatr.* 222, 321–326.
- [188] Feinstein, J.A., Benson, D.W., Dubin, A.M., Cohen, M.S., Maxey, D.M., Mahle, W.T., Pahl, E., Villafañe, J., Bhatt, A.B., Peng, L.F., et al. (2012). Hypoplastic Left Heart Syndrome: Current Considerations and Expectations. *J. Am. Coll. Cardiol.* 59, S1–S42.
- [189] Clark, E.B. (1996). Pathogenetic mechanisms of congenital cardiovascular malformations revisited. *Semin. Perinatol.* 20, 465–472.
- [190] Tchervenkov, C.I., Jacobs, J.P., Weinberg, P.M., Aiello, V.D., Béland, M.J., Colan, S.D., Elliott, M.J., Franklin, R.C.G., Gaynor, J.W., Krogmann, O.N., et al. (2006). The nomenclature, definition and classification of hypoplastic left heart syndrome. *Cardiol. Young* 16, 339–368.

- [191] Jacobs, J.P., O'Brien, S.M., Chai, P.J., Morell, V.O., Lindberg, H.L., and Quintessenza, J.A. (2008). Management of 239 patients with hypoplastic left heart syndrome and related malformations from 1993 to 2007. *Ann. Thorac. Surg.* 85, 1691–1696; discussion 1697.
- [192] Haas, N.A., and Kleideiter, U. (2011). *Kinderkardiologie: Klinik und Praxis der Herzerkrankungen bei Kindern, Jugendlichen und jungen Erwa* (Georg Thieme Verlag).
- [193] Furck, A.K., Uebing, A., Hansen, J.H., Scheewe, J., Jung, O., Fischer, G., Rickers, C., Holland-Letz, T., and Kramer, H.-H. (2010). Outcome of the Norwood operation in patients with hypoplastic left heart syndrome: a 12-year single-center survey. *J. Thorac. Cardiovasc. Surg.* 139, 359–365.
- [194] Hansen, J.H., Furck, A.K., Petko, C., Buchholz-Berdau, R., Voges, I., Scheewe, J., Rickers, C., and Kramer, H.-H. (2012). Use of surveillance criteria reduces interstage mortality after the Norwood operation for hypoplastic left heart syndrome. *Eur. J. Cardio-Thorac. Surg. Off. J. Eur. Assoc. Cardio-Thorac. Surg.* 41, 1013–1018.
- [195] Galli, K.K., Zimmerman, R.A., Jarvik, G.P., Wernovsky, G., Kuypers, M.K., Clancy, R.R., Montenegro, L.M., Mahle, W.T., Newman, M.F., Saunders, A.M., et al. (2004). Periventricular leukomalacia is common after neonatal cardiac surgery. *J. Thorac. Cardiovasc. Surg.* 127, 692–704.
- [196] Iascone, M., Ciccone, R., Galletti, L., Marchetti, D., Seddio, F., Lincesso, A.R., Pezzoli, L., Vetro, A., Barachetti, D., Boni, L., et al. (2012). Identification of de novo mutations and rare variants in hypoplastic left heart syndrome. *Clin. Genet.* 81, 542–554.
- [197] Feit, L.R., Copel, J.A., and Kleinman, C.S. (1991). Foramen ovale size in the normal and abnormal human fetal heart: an indicator of transatrial flow physiology. *Ultrasound Obstet. Gynecol. Off. J. Int. Soc. Ultrasound Obstet. Gynecol.* 1, 313–319.
- [198] Chin, A.J., Weinberg, P.M., and Barber, G. (1990). Subcostal two-dimensional echocardiographic identification of anomalous attachment of septum primum in patients with left atrioventricular valve underdevelopment. *J. Am. Coll. Cardiol.* 15, 678–681.
- [199] Rychik, J., Rome, J.J., Collins, M.H., DeCampi, W.M., and Spray, T.L. (1999). The hypoplastic left heart syndrome with intact atrial septum: atrial morphology, pulmonary vascular histopathology and outcome. *J. Am. Coll. Cardiol.* 34, 554–560.
- [200] Grossfeld, P.D. (1999). The genetics of hypoplastic left heart syndrome. *Cardiol. Young* 9, 627–632.
- [201] Harh, J.Y., Paul, M.H., Gallen, W.J., Friedberg, D.Z., and Kaplan, S. (1973). Experimental production of hypoplastic left heart syndrome in the chick embryo. *Am. J. Cardiol.* 31, 51–56.
- [202] Fishman, N.H., Hof, R.B., Rudolph, A.M., and Heymann, M.A. (1978). Models of congenital heart disease in fetal lambs. *Circulation* 58, 354–364.
- [203] Sedmera, D., Hu, N., Weiss, K.M., Keller, B.B., Denslow, S., and Thompson, R.P. (2002). Cellular changes in experimental left heart hypoplasia. *Anat. Rec.* 267, 137–145.

- [204] Mäkikallio, K., McElhinney, D.B., Levine, J.C., Marx, G.R., Colan, S.D., Marshall, A.C., Lock, J.E., Marcus, E.N., and Tworetzky, W. (2006). Fetal aortic valve stenosis and the evolution of hypoplastic left heart syndrome: patient selection for fetal intervention. *Circulation* 113, 1401–1405.
- [205] Hove, J.R., Köster, R.W., Forouhar, A.S., Acevedo-Bolton, G., Fraser, S.E., and Gharib, M. (2003). Intracardiac fluid forces are an essential epigenetic factor for embryonic cardiogenesis. *Nature* 421, 172–177.
- [206] Groenendijk, B.C.W., Van der Heiden, K., Hierck, B.P., and Poelmann, R.E. (2007). The role of shear stress on ET-1, KLF2, and NOS-3 expression in the developing cardiovascular system of chicken embryos in a venous ligation model. *Physiol. Bethesda Md* 22, 380–389.
- [207] Natowicz, M., Chatten, J., Clancy, R., Conard, K., Glauser, T., Huff, D., Lin, A., Norwood, W., Rorke, L.B., and Uri, A. (1988). Genetic disorders and major extracardiac anomalies associated with the hypoplastic left heart syndrome. *Pediatrics* 82, 698–706.
- [208] Allan, L.D., Sharland, G.K., Milburn, A., Lockhart, S.M., Groves, A.M., Anderson, R.H., Cook, A.C., and Fagg, N.L. (1994). Prospective diagnosis of 1,006 consecutive cases of congenital heart disease in the fetus. *J. Am. Coll. Cardiol.* 23, 1452–1458.
- [209] Raymond, F.L., Simpson, J.M., Sharland, G.K., and Ogilvie Mackie, C.M. (1997). Fetal echocardiography as a predictor of chromosomal abnormality. *Lancet Lond. Engl.* 350, 930.
- [210] Patel, A., Hickey, E., Mavroudis, C., Jacobs, J.P., Jacobs, M.L., Backer, C.L., Gevitz, M., and Mavroudis, C.D. (2010). Impact of noncardiac congenital and genetic abnormalities on outcomes in hypoplastic left heart syndrome. *Ann. Thorac. Surg.* 89, 1805–1813; discussion 1813–1814.
- [211] Natowicz, M., and Kelley, R.I. (1987). Association of Turner syndrome with hypoplastic left-heart syndrome. *Am. J. Dis. Child.* 1960 141, 218–220.
- [212] Antonelli, D., Antonelli, J., and Rosenfeld, T. (1990). Noonan's syndrome associated with hypoplastic left heart. *Cardiology* 77, 62–65.
- [213] Glauser, T.A., Zackai, E., Weinberg, P., and Clancy, R. (1989). Holt-Oram syndrome associated with the hypoplastic left heart syndrome. *Clin. Genet.* 36, 69–72.
- [214] Mattina, T., Perrotta, C.S., and Grossfeld, P. (2009). Jacobsen syndrome. *Orphanet J. Rare Dis.* 4, 9.
- [215] Sanchez-Valle, A., Pierpont, M.E., and Potocki, L. (2011). The severe end of the spectrum: Hypoplastic left heart in Potocki-Lupski syndrome. *Am. J. Med. Genet. A.* 155A, 363–366.
- [216] Kung, G.C., Chang, P.M., Sklansky, M.S., and Randolph, L.M. (2010). Hypoplastic left heart syndrome in patients with Kabuki syndrome. *Pediatr. Cardiol.* 31, 138–141.
- [217] Campbell, C.L., Collins, R.T., and Zarate, Y.A. (2014). Severe neonatal presentation of Kleeftstra syndrome in a patient with hypoplastic left heart syndrome and 9q34.3 microdeletion. *Birt. Defects Res. A. Clin. Mol. Teratol.* 100, 985–990.

- [218] Gerboni, S., Sabatino, G., Mingarelli, R., and Dallapiccola, B. (1993). Coarctation of the aorta, interrupted aortic arch, and hypoplastic left heart syndrome in three generations. *J. Med. Genet.* 30, 328–329.
- [219] Grossfeld, P.D., Mattina, T., Lai, Z., Favier, R., Jones, K.L., Cotter, F., and Jones, C. (2004). The 11q terminal deletion disorder: a prospective study of 110 cases. *Am. J. Med. Genet. A.* 129A, 51–61.
- [220] Grossfeld, P.D. (2007). Hypoplastic left heart syndrome: it is all in the genes. *J. Am. Coll. Cardiol.* 50, 1596–1597.
- [221] Stallmeyer, B., Fenge, H., Nowak-Göttl, U., and Schulze-Bahr, E. (2010). Mutational spectrum in the cardiac transcription factor gene NKX2.5 (CSX) associated with congenital heart disease. *Clin. Genet.* 78, 533–540.
- [222] Dasgupta, C., Martinez, A.M., Zuppan, C.W., Shah, M.M., Bailey, L.L., and Fletcher, W.H. (2001). Identification of connexin43 (alpha1) gap junction gene mutations in patients with hypoplastic left heart syndrome by denaturing gradient gel electrophoresis (DGGE). *Mutat. Res.* 479, 173–186.
- [223] Timmerman, L.A., Grego-Bessa, J., Raya, A., Bertrán, E., Pérez-Pomares, J.M., Díez, J., Aranda, S., Palomo, S., McCormick, F., Izpisua-Belmonte, J.C., et al. (2004). Notch promotes epithelial-mesenchymal transition during cardiac development and oncogenic transformation. *Genes Dev.* 18, 99–115.
- [224] Luna-Zurita, L., Prados, B., Grego-Bessa, J., Luxán, G., del Monte, G., Benguría, A., Adams, R.H., Pérez-Pomares, J.M., and de la Pompa, J.L. (2010). Integration of a Notch-dependent mesenchymal gene program and Bmp2-driven cell invasiveness regulates murine cardiac valve formation. *J. Clin. Invest.* 120, 3493–3507.
- [225] McBride, K.L., Riley, M.F., Zender, G.A., Fitzgerald-Butt, S.M., Towbin, J.A., Belmont, J.W., and Cole, S.E. (2008). NOTCH1 mutations in individuals with left ventricular outflow tract malformations reduce ligand-induced signaling. *Hum. Mol. Genet.* 17, 2886–2893.
- [226] Theis, J.L., Hrstka, S.C.L., Evans, J.M., O’Byrne, M.M., de Andrade, M., O’Leary, P.W., Nelson, T.J., and Olson, T.M. (2015). Compound heterozygous NOTCH1 mutations underlie impaired cardiogenesis in a patient with hypoplastic left heart syndrome. *Hum. Genet.*
- [227] Liu, X., Saeed, S., Li, Y., Francis, R., Tobita, K., Kim, A.J., Shung, W., Leatherbury, L., and Lo, C.W. (2012). Abstract 19570: A Multigenic Etiology Of Hypoplastic Left Heart Syndrome: An Analysis Based on Three Novel Mutant Mouse Models of Hypoplastic Left Heart Syndrome. *Circulation* 126, A19570.
- [228] Hinton, R.B., Martin, L.J., Rame-Gowda, S., Tabangin, M.E., Cripe, L.H., and Benson, D.W. (2009). Hypoplastic left heart syndrome links to chromosomes 10q and 6q and is genetically related to bicuspid aortic valve. *J. Am. Coll. Cardiol.* 53, 1065–1071.

- [229] McBride, K.L., Zender, G.A., Fitzgerald-Butt, S.M., Koehler, D., Menesses-Diaz, A., Fernbach, S., Lee, K., Towbin, J.A., Leal, S., and Belmont, J.W. (2009). Linkage analysis of left ventricular outflow tract malformations (aortic valve stenosis, coarctation of the aorta, and hypoplastic left heart syndrome). *Eur. J. Hum. Genet.* 17, 811–819.
- [230] Stankiewicz, P., Sen, P., Bhatt, S.S., Storer, M., Xia, Z., Bejjani, B.A., Ou, Z., Wiszniewska, J., Driscoll, D.J., Maisenbacher, M.K., et al. (2009). Genomic and genic deletions of the FOX gene cluster on 16q24.1 and inactivating mutations of FOXF1 cause alveolar capillary dysplasia and other malformations. *Am. J. Hum. Genet.* 84, 780–791.
- [231] Ferencz, C., Rubin, J.D., McCarter, R.J., Brenner, J.I., Neill, C.A., Perry, L.W., Hepner, S.I., and Downing, J.W. (1985). Congenital heart disease: prevalence at livebirth. The Baltimore-Washington Infant Study. *Am. J. Epidemiol.* 121, 31–36.
- [232] Brenner, J.I., Berg, K.A., Schneider, D.S., Clark, E.B., and Boughman, J.A. (1989). Cardiac malformations in relatives of infants with hypoplastic left-heart syndrome. *Am. J. Dis. Child.* 1960 143, 1492–1494.
- [233] Loffredo, C.A., Chokkalingam, A., Sill, A.M., Boughman, J.A., Clark, E.B., Scheel, J., and Brenner, J.I. (2004). Prevalence of congenital cardiovascular malformations among relatives of infants with hypoplastic left heart, coarctation of the aorta, and d-transposition of the great arteries. *Am. J. Med. Genet. A.* 124A, 225–230.
- [234] Hinton Jr, R.B., Martin, L.J., Tabangin, M.E., Mazwi, M.L., Cripe, L.H., and Benson, D.W. (2007). Hypoplastic Left Heart Syndrome Is Heritable. *J. Am. Coll. Cardiol.* 50, 1590–1595.
- [235] Kelle, A.M., Qureshi, M.Y., Olson, T.M., Eidem, B.W., and O’Leary, P.W. (2015). Familial Incidence of Cardiovascular Malformations in Hypoplastic Left Heart Syndrome. *Am. J. Cardiol.*
- [236] Huntington, K., Hunter, A.G., and Chan, K.L. (1997). A prospective study to assess the frequency of familial clustering of congenital bicuspid aortic valve. *J. Am. Coll. Cardiol.* 30, 1809–1812.
- [237] McBride, K.L., Pignatelli, R., Lewin, M., Ho, T., Fernbach, S., Menesses, A., Lam, W., Leal, S.M., Kaplan, N., Schliekelman, P., et al. (2005). Inheritance Analysis of Congenital Left Ventricular Outflow Tract Obstruction Malformations: Segregation, Multiplex Relative Risk, and Heritability. *Am. J. Med. Genet. A.* 134, 180–186.
- [238] Hales, A.R., and Mahle, W.T. (2014). Echocardiography screening of siblings of children with bicuspid aortic valve. *Pediatrics* 133, e1212–e1217.
- [239] Grossfeld, P., Ye, M., and Harvey, R. (2009). Hypoplastic left heart syndrome: new genetic insights. *J. Am. Coll. Cardiol.* 53, 1072–1074.
- [240] Swain, S. (2011). Evaluation of Recurrence Risks for Left-Sided Cardiac Lesions. UT GSBS Diss. Theses Open Access.
- [241] Kobayashi, J., Yoshida, M., Tarui, S., Hirata, M., Nagai, Y., Kasahara, S., Naruse, K., Ito, H., Sano, S., and Oh, H. (2014). Directed Differentiation of Patient-Specific Induced Pluripotent Stem Cells Identifies the Transcriptional Repression and Epigenetic Modification of NKX2-5, HAND1, and NOTCH1 in Hypoplastic Left Heart Syndrome. *PLoS ONE* 9, e102796.

- [242] Campos, E.I., and Reinberg, D. (2009). Histones: annotating chromatin. *Annu. Rev. Genet.* 43, 559–599.
- [243] Bernstein, B.E., Mikkelsen, T.S., Xie, X., Kamal, M., Huebert, D.J., Cuff, J., Fry, B., Meissner, A., Wernig, M., Plath, K., et al. (2006). A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* 125, 315–326.
- [244] Pinskaya, M., and Morillon, A. (2009). Histone H3 lysine 4 di-methylation: a novel mark for transcriptional fidelity? *Epigenetics* 4, 302–306.
- [245] Paige, S.L., Thomas, S., Stoick-Cooper, C.L., Wang, H., Maves, L., Sandstrom, R., Pabon, L., Reinecke, H., Pratt, G., Keller, G., et al. (2012). A temporal chromatin signature in human embryonic stem cells identifies regulators of cardiac development. *Cell* 151, 221–232.
- [246] Wessels, M.W., Berger, R.M.F., Frohn-Mulder, I.M.E., Roos-Hesselink, J.W., Hoogeboom, J.J.M., Mancini, G.S., Bartelings, M.M., Krijger, R. de, Wladimiroff, J.W., Niermeijer, M.F., et al. (2005). Autosomal dominant inheritance of left ventricular outflow tract obstruction. *Am. J. Med. Genet. A.* 134A, 171–179.
- [247] Arndt, A.-K., Petko, C., Scheewe, J., and Kramer, H.-H. (2012). Staged surgical palliation in identical twins with concordance for hypoplastic left heart syndrome (HLHS). *Klin. Pädiatr.* 224, 268–269.
- [248] Zhu, X. (2014). *Surgical Atlas of Cardiac Anatomy* (Springer).
- [249] Schwegler, J.S. (2011). *Der Mensch - Anatomie und Physiologie* (Georg Thieme Verlag).
- [250] Pulvertaft, J.V. (1964). CYTOLOGY OF BURKITT'S TUMOUR (AFRICAN LYMPHOMA). *Lancet* 1, 238–240.
- [251] Kent, W.J., Sugnet, C.W., Furey, T.S., Roskin, K.M., Pringle, T.H., Zahler, A.M., and Haussler, D. (2002). The human genome browser at UCSC. *Genome Res.* 12, 996–1006.
- [252] Huber, W., Carey, V.J., Gentleman, R., Anders, S., Carlson, M., Carvalho, B.S., Bravo, H.C., Davis, S., Gatto, L., Girke, T., et al. (2015). Orchestrating high-throughput genomic analysis with Bioconductor. *Nat. Methods* 12, 115–121.
- [253] Pages, H., Aboyoun, P., Gentleman, R., and DebRoy, S. Biostrings: String objects representing biological sequences, and matching algorithms.
- [254] Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26, 139–140.
- [255] Thompson, R.F., Suzuki, M., Lau, K.W., and Grealley, J.M. (2009). A pipeline for the quantitative analysis of CG dinucleotide methylation using mass spectrometry. *Bioinformatics* 25, 2164–2170.
- [256] Assenov, Y., Müller, F., Lutsik, P., Walter, J., Lengauer, T., and Bock, C. (2014). Comprehensive analysis of DNA methylation data with RnBeads. *Nat. Methods* 11, 1138–1140.
- [257] Consortium, T. 1000 G.P. (2010). A map of human genome variation from population-scale sequencing. *Nature* 467, 1061–1073.
- [258] Kent, W.J. (2002). BLAT--the BLAST-like alignment tool. *Genome Res.* 12, 656–664.

- [259] Sherry, S.T., Ward, M.-H., Kholodov, M., Baker, J., Phan, L., Smigielski, E.M., and Sirotkin, K. (2001). dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res.* 29, 308–311.
- [260] MacDonald, J.R., Ziman, R., Yuen, R.K.C., Feuk, L., and Scherer, S.W. (2014). The Database of Genomic Variants: a curated collection of structural variation in the human genome. *Nucleic Acids Res.* 42, D986–D992.
- [261] Dinkel, H., Michael, S., Weatheritt, R.J., Davey, N.E., Roey, K.V., Altenberg, B., Toedt, G., Uyar, B., Seiler, M., Budd, A., et al. (2012). ELM—the database of eukaryotic linear motifs. *Nucleic Acids Res.* 40, D242–D251.
- [262] Rebhan, M., Chalifa-Caspi, V., Prilusky, J., and Lancet, D. (1998). GeneCards: a novel functional genomics compendium with automated data mining and query reformulation support. *Bioinforma. Oxf. Engl.* 14, 656–664.
- [263] Warde-Farley, D., Donaldson, S.L., Comes, O., Zuberi, K., Badrawi, R., Chao, P., Franz, M., Grouios, C., Kazi, F., Lopes, C.T., et al. (2010). The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. *Nucleic Acids Res.* 38, W214–W220.
- [264] Petrovski, S., Wang, Q., Heinzen, E.L., Allen, A.S., and Goldstein, D.B. (2013). Genic Intolerance to Functional Variation and the Interpretation of Personal Genomes. *PLoS Genet* 9, e1003709.
- [265] Wain, H.M., Bruford, E.A., Lovering, R.C., Lush, M.J., Wright, M.W., and Povey, S. (2002). Guidelines for human gene nomenclature. *Genomics* 79, 464–470.
- [266] den Dunnen, J.T., and Antonarakis, S.E. (2000). Mutation nomenclature extensions and suggestions to describe complex mutations: a discussion. *Hum. Mutat.* 15, 7–12.
- [267] Shaw, D.R. (2009). Searching the Mouse Genome Informatics (MGI) resources for information on mouse biology from genotype to phenotype. *Curr. Protoc. Bioinforma.* Ed. Board Andreas Baxevanis AI Chapter 1, Unit1.7.
- [268] Schwarz, J.M., Rödelsperger, C., Schuelke, M., and Seelow, D. (2010). MutationTaster evaluates disease-causing potential of sequence alterations. *Nat. Methods* 7, 575–576.
- [269] Kibbe, W.A. (2007). OligoCalc: an online oligonucleotide properties calculator. *Nucleic Acids Res.* 35, W43–W46.
- [270] Hamosh, A., Scott, A.F., Amberger, J.S., Bocchini, C.A., and McKusick, V.A. (2005). Online Mendelian Inheritance in Man (OMIM), a knowledgebase of human genes and genetic disorders. *Nucleic Acids Res.* 33, D514–D517.
- [271] Schaefer, C.F., Anthony, K., Krupa, S., Buchoff, J., Day, M., Hannay, T., and Buetow, K.H. (2009). PID: the Pathway Interaction Database. *Nucleic Acids Res.* 37, D674–D679.
- [272] Adzhubei, I., Jordan, D.M., and Sunyaev, S.R. (2013). Predicting functional effect of human missense mutations using PolyPhen-2. *Curr. Protoc. Hum. Genet.* Editor. Board Jonathan Haines AI Chapter 7, Unit7.20.
- [273] Haw, R., Hermjakob, H., D'Eustachio, P., and Stein, L. (2011). Reactome pathway analysis to enrich biological discovery in proteomics data sets. *Proteomics* 11, 3598–3613.

- [274] Franceschini, A., Szklarczyk, D., Frankild, S., Kuhn, M., Simonovic, M., Roth, A., Lin, J., Minguez, P., Bork, P., von Mering, C., et al. (2013). STRING v9.1: protein-protein interaction networks, with increased coverage and integration. *Nucleic Acids Res.* 41, D808–D815.
- [275] Gerstein, M.B., Kundaje, A., Hariharan, M., Landt, S.G., Yan, K.-K., Cheng, C., Mu, X.J., Khurana, E., Rozowsky, J., Alexander, R., et al. (2012). Architecture of the human regulatory network derived from ENCODE data. *Nature* 489, 91–100.
- [276] Wang, J., Zhuang, J., Iyer, S., Lin, X., Whitfield, T.W., Greven, M.C., Pierce, B.G., Dong, X., Kundaje, A., Cheng, Y., et al. (2012). Sequence features and chromatin structure around the genomic regions bound by 119 human transcription factors. *Genome Res.* 22, 1798–1812.
- [277] Wang, J., Zhuang, J., Iyer, S., Lin, X.-Y., Greven, M.C., Kim, B.-H., Moore, J., Pierce, B.G., Dong, X., Virgil, D., et al. (2013a). Factorbook.org: a Wiki-based database for transcription factor-binding data generated by the ENCODE consortium. *Nucleic Acids Res.* 41, D171–D176.
- [278] Wang, J., Duncan, D., Shi, Z., and Zhang, B. (2013b). WEB-based GENE SeT ANALYSIS Toolkit (WebGestalt): update 2013. *Nucleic Acids Res.* 41, W77–W83.
- [279] Sah, S., Chen, L., Houghton, J., Kempainen, J., Marko, A.C., Zeigler, R., and Latham, G.J. (2013). Functional DNA quantification guides accurate next-generation sequencing mutation detection in formalin-fixed, paraffin-embedded tumor biopsies. *Genome Med.* 5, 77.
- [280] Sanger, F., Nicklen, S., and Coulson, A.R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. U. S. A.* 74, 5463–5467.
- [281] Lottspeich, F., and Engels, J.W. (2008). *Bioanalytik* (Spektrum Akademischer Verlag).
- [282] Metzker, M.L. (2010). Sequencing technologies — the next generation. *Nat. Rev. Genet.* 11, 31–46.
- [283] Hitz, Marc-Phillip (Mai 2015). Genetic determinants in left ventricular outflow tract obstruction (LVOTO). Noch nicht veröffentlichte Dissertation, Mai 2015. Darwin College University of Cambridge.
- [284] Hayatsu, H., Wataya, Y., Kai, K., and Iida, S. (1970). Reaction of sodium bisulfite with uracil, cytosine, and their derivatives. *Biochemistry (Mosc.)* 9, 2858–2865.
- [285] Shapiro, R., Servis, R.E., and Welcher, M. (1970). Reactions of Uracil and Cytosine Derivatives with Sodium Bisulfite. *J. Am. Chem. Soc.* 92, 422–424.
- [286] Bibikova, M., Barnes, B., Tsan, C., Ho, V., Klotzle, B., Le, J.M., Delano, D., Zhang, L., Schroth, G.P., Gunderson, K.L., et al. (2011). High density DNA methylation array with single CpG site resolution. *Genomics* 98, 288–295.
- [287] Maksimovic, J., Gordon, L., and Oshlack, A. (2012). SWAN: Subset-quantile Within Array Normalization for Illumina Infinium HumanMethylation450 BeadChips. *Genome Biol.* 13, R44.
- [288] Chen, Y., Lemire, M., Choufani, S., Butcher, D.T., Grafodatskaya, D., Zanke, B.W., Gallinger, S., Hudson, T.J., and Weksberg, R. (2013). Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. *Epigenetics Off. J. DNA Methylation Soc.* 8, 203–209.

- [289] Dupont, J.-M., Tost, J., Jammes, H., and Gut, I.G. (2004). De novo quantitative bisulfite sequencing using the pyrosequencing technology. *Anal. Biochem.* 333, 119–127.
- [290] Tost, J., and Gut, I.G. (2007). DNA methylation analysis by pyrosequencing. *Nat. Protoc.* 2, 2265–2275.
- [291] Ehrich, M., Nelson, M.R., Stanssens, P., Zabeau, M., Liloglou, T., Xinarianos, G., Cantor, C.R., Field, J.K., and van den Boom, D. (2005). Quantitative high-throughput analysis of DNA methylation patterns by base-specific cleavage and mass spectrometry. *Proc. Natl. Acad. Sci. U. S. A.* 102, 15785–15790.
- [292] Wang, Z., Gerstein, M., and Snyder, M. (2009). RNA-Seq: a revolutionary tool for transcriptomics. *Nat. Rev. Genet.* 10, 57–63.
- [293] Martin M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnetjournal, North America.* 17
- [294] Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S.L. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* 14, R36.
- [295] Bustin, S.A., Benes, V., Garson, J.A., Hellemans, J., Huggett, J., Kubista, M., Mueller, R., Nolan, T., Pfaffl, M.W., Shipley, G.L., et al. (2009). The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clin. Chem.* 55, 611–622.
- [296] Du, P., Zhang, X., Huang, C.-C., Jafari, N., Kibbe, W.A., Hou, L., and Lin, S.M. (2010). Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis. *BMC Bioinformatics* 11, 587.
- [297] Benjamini, Y., and Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. R. Stat. Soc. Ser. B Methodol.* 57, 289–300.
- [298] Auer, P.L., and Doerge, R.W. (2010). Statistical Design and Analysis of RNA Sequencing Data. *Genetics* 185, 405–416.
- [299] Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T., et al. (2000). Gene Ontology: tool for the unification of biology. *Nat. Genet.* 25, 25–29.
- [300] Whittemore, R., Hobbins, J.C., and Engle, M.A. (1982). Pregnancy and its outcome in women with and without surgical treatment of congenital heart disease. *Am. J. Cardiol.* 50, 641–651.
- [301] Nabulsi, M.M., Tamim, H., Sabbagh, M., Obeid, M.Y., Yunis, K.A., and Bitar, F.F. (2003). Parental consanguinity and congenital heart malformations in a developing country. *Am. J. Med. Genet. A.* 116A, 342–347.
- [302] Yunis, K., Khalid, Y., Mumtaz, G., Ghina, M., Bitar, F., Fadi, B., Chamseddine, F., Fadi, C., Kassab, M., May, K., et al. (2006). Consanguineous marriage and congenital heart defects: a case-control study in the neonatal period. *Am. J. Med. Genet. A.* 140, 1524–1530.

- [303] Al Turki, S., Manickaraj, A.K., Mercer, C.L., Gerety, S.S., Hitz, M.-P., Lindsay, S., D'Alessandro, L.C.A., Swaminathan, G.J., Bentham, J., Arndt, A.-K., et al. (2014). Rare variants in NR2F2 cause congenital heart defects in humans. *Am. J. Hum. Genet.* 94, 574–585.
- [304] Emanuel, R., Withers, R., O'Brien, K., Ross, P., and Feizi, O. (1978). Congenitally bicuspid aortic valves. Clinicogenetic study of 41 families. *Br. Heart J.* 40, 1402–1407.
- [305] Mendel, G. (1866). *Versuche über Pflanzen-Hybriden* (Brünn: Naturforschender Verein).
- [306] Jorgensen, M., McPherson, E., Zaleski, C., Shivaram, P., and Cold, C. (2014). Stillbirth: The heart of the matter. *Am. J. Med. Genet. A.* 164, 691–699.
- [307] Thorsson, T., Russell, W.W., El-Kashlan, N., Soemedi, R., Levine, J., Geisler, S.B., Ackley, T., Tomita-Mitchell, A., Rosenfeld, J.A., Töpf, A., et al. (2015). Chromosomal Imbalances in Patients with Congenital Cardiac Defects: A Meta-analysis Reveals Novel Potential Critical Regions Involved in Heart Development. *Congenit. Heart Dis.* 10, 193–208.
- [308] Clement, C.A., Kristensen, S.G., Møllgård, K., Pazour, G.J., Yoder, B.K., Larsen, L.A., and Christensen, S.T. (2009). The primary cilium coordinates early cardiogenesis and hedgehog signaling in cardiomyocyte differentiation. *J. Cell Sci.* 122, 3070–3082.
- [309] Riley, P., Anson-Cartwright, L., and Cross, J.C. (1998). The Hand1 bHLH transcription factor is essential for placentation and cardiac morphogenesis. *Nat. Genet.* 18, 271–275.
- [310] Firulli, A.B., McFadden, D.G., Lin, Q., Srivastava, D., and Olson, E.N. (1998). Heart and extra-embryonic mesodermal defects in mouse embryos lacking the bHLH transcription factor Hand1. *Nat. Genet.* 18, 266–270.
- [311] Kulis, M., and Esteller, M. (2010). DNA methylation and cancer. *Adv. Genet.* 70, 27–56.
- [312] Hansen, K.D., Timp, W., Bravo, H.C., Sabunciyan, S., Langmead, B., McDonald, O.G., Wen, B., Wu, H., Liu, Y., Diep, D., et al. (2011). Increased methylation variation in epigenetic domains across cancer types. *Nat. Genet.* 43, 768–775.
- [313] Fradin, D., Le Fur, S., Mille, C., Naoui, N., Groves, C., Zelenika, D., McCarthy, M.I., Lathrop, M., and Bougnères, P. (2012). Association of the CpG methylation pattern of the proximal insulin gene promoter with type 1 diabetes. *PLoS One* 7, e36278.
- [314] Stefan, M., Zhang, W., Concepcion, E., Yi, Z., and Tomer, Y. (2014). DNA methylation profiles in type 1 diabetes twins point to strong epigenetic effects on etiology. *J. Autoimmun.* 50, 33–37.
- [315] Connor, C.M., and Akbarian, S. (2008). DNA methylation changes in schizophrenia and bipolar disorder. *Epigenetics Off. J. DNA Methylation Soc.* 3, 55–58.
- [316] Rao, J.S., Keleshian, V.L., Klein, S., and Rapoport, S.I. (2012). Epigenetic modifications in frontal cortex from Alzheimer's disease and bipolar disorder patients. *Transl. Psychiatry* 2, e132.
- [317] Ohtani, K., and Dimmeler, S. (2011). Epigenetic regulation of cardiovascular differentiation. *Cardiovasc. Res.* cvr019.
- [318] Wamstad, J.A., Alexander, J.M., Truty, R.M., Shrikumar, A., Li, F., Eilertson, K.E., Ding, H., Wylie, J.N., Pico, A.R., Capra, J.A., et al. (2012). Dynamic and coordinated epigenetic regulation of developmental transitions in the cardiac lineage. *Cell* 151, 206–220.

- [319] Sim, C.B., Ziemann, M., Kaspi, A., Harikrishnan, K.N., Ooi, J., Khurana, I., Chang, L., Hudson, J.E., El-Osta, A., and Porrello, E.R. (2014). Dynamic changes in the cardiac methylome during postnatal development. *FASEB J.* fj.14–264093.
- [320] Erbel, R., and Eggebrecht, H. (2006). Aortic dimensions and the risk of dissection. *Heart* 92, 137–142.
- [321] Arndt, A.-K., Schafer, S., Drenckhahn, J.-D., Sabeh, M.K., Plovie, E.R., Caliebe, A., Klopocki, E., Musso, G., Werdich, A.A., Kalwa, H., et al. (2013). Fine mapping of the 1p36 deletion syndrome identifies mutation of PRDM16 as a cause of cardiomyopathy. *Am. J. Hum. Genet.* 93, 67–77.
- [322] Slaney, S.F., Oldridge, M., Hurst, J.A., Moriss-Kay, G.M., Hall, C.M., Poole, M.D., and Wilkie, A.O. (1996). Differential effects of FGFR2 mutations on syndactyly and cleft palate in Apert syndrome. *Am. J. Hum. Genet.* 58, 923–932.
- [323] Ewart, A.K., Jin, W., Atkinson, D., Morris, C.A., and Keating, M.T. (1994). Supravalvular aortic stenosis associated with a deletion disrupting the elastin gene. *J. Clin. Invest.* 93, 1071–1077.
- [324] Urbán, Z., Michels, V.V., Thibodeau, S.N., Davis, E.C., Bonnefont, J.P., Munnich, A., Eyskens, B., Gewillig, M., Devriendt, K., and Boyd, C.D. (2000). Isolated supravalvular aortic stenosis: functional haploinsufficiency of the elastin gene as a result of nonsense-mediated decay. *Hum. Genet.* 106, 577–588.
- [325] Doyle, A.J., Doyle, J.J., Bessling, S.L., Maragh, S., Lindsay, M.E., Schepers, D., Gillis, E., Mortier, G., Homfray, T., Sauls, K., et al. (2012). Mutations in the TGF- β repressor SKI cause Shprintzen-Goldberg syndrome with aortic aneurysm. *Nat. Genet.* 44, 1249–1254.
- [326] Pasutto, F., Sticht, H., Hammersen, G., Gillessen-Kaesbach, G., Fitzpatrick, D.R., Nürnberg, G., Brasch, F., Schirmer-Zimmermann, H., Tolmie, J.L., Chitayat, D., et al. (2007). Mutations in STRA6 cause a broad spectrum of malformations including anophthalmia, congenital heart defects, diaphragmatic hernia, alveolar capillary dysplasia, lung hypoplasia, and mental retardation. *Am. J. Hum. Genet.* 80, 550–560.
- [327] Chassaing, N., Ragge, N., Kariminejad, A., Buffet, A., Ghaderi-Sohi, S., Martinovic, J., and Calvas, P. (2013). Mutation analysis of the STRA6 gene in isolated and non-isolated anophthalmia/microphthalmia. *Clin. Genet.* 83, 244–250.
- [328] Basson, C.T., Huang, T., Lin, R.C., Bachinsky, D.R., Weremowicz, S., Vaglio, A., Bruzzone, R., Quadrelli, R., Lerone, M., Romeo, G., et al. (1999). Different TBX5 interactions in heart and limb defined by Holt-Oram syndrome mutations. *Proc. Natl. Acad. Sci. U. S. A.* 96, 2919–2924.
- [329] Mori, A.D., and Bruneau, B.G. (2004). TBX5 mutations and congenital heart disease: Holt-Oram syndrome revealed. *Curr. Opin. Cardiol.* 19, 211–215.
- [330] Amiel, J., Espinosa-Parrilla, Y., Steffann, J., Gosset, P., Pelet, A., Prieur, M., Boute, O., Choiset, A., Lacombe, D., Philip, N., et al. (2001). Large-scale deletions and SMADIP1 truncating mutations in syndromic Hirschsprung disease with involvement of midline structures. *Am. J. Hum. Genet.* 69, 1370–1377.

- [331] Ishihara, N., Yamada, K., Yamada, Y., Miura, K., Kato, J., Kuwabara, N., Hara, Y., Kobayashi, Y., Hoshino, K., Nomura, Y., et al. (2004). Clinical and molecular analysis of Mowat-Wilson syndrome associated with ZFX1B mutations and deletions at 2q22-q24.1. *J. Med. Genet.* 41, 387–393.
- [332] Cheng, Z., Wang, J., Su, D., Pan, H., Huang, G., Li, X., Li, Z., Shen, A., Xie, X., Wang, B., et al. (2011). Two novel mutations of the IRX4 gene in patients with congenital heart disease. *Hum. Genet.* 130, 657–662.
- [333] Furey, T.S. (2012). ChIP-seq and beyond: new and improved methodologies to detect and characterize protein-DNA interactions. *Nat. Rev. Genet.* 13, 840–852.
- [334] Li, X., Nair, A., Wang, S., and Wang, L. (2015). Quality Control of RNA-Seq Experiments. In *RNA Bioinformatics*, E. Picardi, ed. (Springer New York), pp. 137–146.
- [335] Illingworth, R.S., and Bird, A.P. (2009). CpG islands – “A rough guide.” *FEBS Lett.* 583, 1713–1720.
- [336] Weinberg, P.M., Chin, A.J., Murphy, J.D., Pigott, J.D., and Norwood, W.I. (1986). Postmortem echocardiography and tomographic anatomy of hypoplastic left heart syndrome after palliative surgery. *Am. J. Cardiol.* 58, 1228–1232.
- [337] Porrello, E.R., Mahmoud, A.I., Simpson, E., Hill, J.A., Richardson, J.A., Olson, E.N., and Sadek, H.A. (2011). Transient regenerative potential of the neonatal mouse heart. *Science* 331, 1078–1080.
- [338] Porrello, E.R., Mahmoud, A.I., Simpson, E., Johnson, B.A., Grinsfelder, D., Canseco, D., Mammen, P.P., Rothermel, B.A., Olson, E.N., and Sadek, H.A. (2013). Regulation of neonatal and adult mammalian heart regeneration by the miR-15 family. *Proc. Natl. Acad. Sci. U. S. A.* 110, 187–192.
- [339] Porrello, E.R., and Olson, E.N. (2014). A neonatal blueprint for cardiac regeneration. *Stem Cell Res.* 13, 556–570.
- [340] Rhein, M., Hagemeyer, L., Klitschar, M., Muschler, M., Bleich, S., and Frieling, H. (2015). DNA methylation results depend on DNA integrity—role of post mortem interval. *Front. Genet.* 6.
- [341] Barrachina, M., and Ferrer, I. (2009). DNA methylation of Alzheimer disease and tauopathy-related genes in postmortem brain. *J. Neuropathol. Exp. Neurol.* 68, 880–891.
- [342] Margulies, M., Egholm, M., Altman, W.E., Attiya, S., Bader, J.S., Bemben, L.A., Berka, J., Braverman, M.S., Chen, Y.-J., Chen, Z., et al. (2005). Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 437, 376–380.
- [343] Shendure, J., Porreca, G.J., Reppas, N.B., Lin, X., McCutcheon, J.P., Rosenbaum, A.M., Wang, M.D., Zhang, K., Mitra, R.D., and Church, G.M. (2005). Accurate multiplex polony sequencing of an evolved bacterial genome. *Science* 309, 1728–1732.
- [344] Hutchison, C.A. (2007). DNA sequencing: bench to bedside and beyond. *Nucleic Acids Res.* 35, 6227–6237.
- [345] Schuster, S.C. (2008). Next-generation sequencing transforms today’s biology. *Nat. Methods* 5, 16–18.

- [346] Consortium, T.E.P. (2012). An integrated encyclopedia of DNA elements in the human genome. *Nature* 489, 57–74.
- [347] Ng, S.B., Turner, E.H., Robertson, P.D., Flygare, S.D., Bigham, A.W., Lee, C., Shaffer, T., Wong, M., Bhattacharjee, A., Eichler, E.E., et al. (2009b). Targeted Capture and Massively Parallel Sequencing of Twelve Human Exomes. *Nature* 461, 272–276.
- [348] The UK10K Consortium (2015). The UK10K project identifies rare variants in health and disease. *Nature* 526, 82–90.
- [349] Gilissen, C., Hoischen, A., Brunner, H.G., and Veltman, J.A. (2011). Unlocking Mendelian disease using exome sequencing. *Genome Biol.* 12, 228.
- [350] Bamshad, M.J., Ng, S.B., Bigham, A.W., Tabor, H.K., Emond, M.J., Nickerson, D.A., and Shendure, J. (2011). Exome sequencing as a tool for Mendelian disease gene discovery. *Nat. Rev. Genet.* 12, 745–755.
- [351] Clark, M.J., Chen, R., Lam, H.Y.K., Karczewski, K.J., Chen, R., Euskirchen, G., Butte, A.J., and Snyder, M. (2011). Performance comparison of exome DNA sequencing technologies. *Nat. Biotechnol.* 29, 908–914.
- [352] Bodi, K., Perera, A.G., Adams, P.S., Bintzler, D., Dewar, K., Grove, D.S., Kieleczawa, J., Lyons, R.H., Neubert, T.A., Noll, A.C., et al. (2013). Comparison of Commercially Available Target Enrichment Methods for Next-Generation Sequencing. *J. Biomol. Tech. JBT* 24, 73–86.
- [353] Chilamakuri, C.S.R., Lorenz, S., Madoui, M.-A., Vodák, D., Sun, J., Hovig, E., Myklebost, O., and Meza-Zepeda, L.A. (2014). Performance comparison of four exome capture systems for deep sequencing. *BMC Genomics* 15, 449.
- [354] Hsu, F., Kent, W.J., Clawson, H., Kuhn, R.M., Diekhans, M., and Haussler, D. (2006). The UCSC Known Genes. *Bioinforma. Oxf. Engl.* 22, 1036–1046.
- [355] Pruitt, K.D., Tatusova, T., Klimke, W., and Maglott, D.R. (2009). NCBI Reference Sequences: current status, policy and new initiatives. *Nucleic Acids Res.* 37, D32–D36.
- [356] Flicek, P., Amode, M.R., Barrell, D., Beal, K., Brent, S., Chen, Y., Clapham, P., Coates, G., Fairley, S., Fitzgerald, S., et al. (2011). Ensembl 2011. *Nucleic Acids Res.* 39, D800–D806.
- [357] Schmitt, M.W., Kennedy, S.R., Salk, J.J., Fox, E.J., Hiatt, J.B., and Loeb, L.A. (2012). Detection of ultra-rare mutations by next-generation sequencing. *Proc. Natl. Acad. Sci. U. S. A.* 109, 14508–14513.
- [358] Miyatake, S., Koshimizu, E., Hayashi, Y.K., Miya, K., Shiina, M., Nakashima, M., Tsurusaki, Y., Miyake, N., Saito, H., Ogata, K., et al. (2014). Deep sequencing detects very-low-grade somatic mosaicism in the unaffected mother of siblings with nemaline myopathy. *Neuromuscul. Disord.* 24, 642–647.
- [359] Martincorena, I., Roshan, A., Gerstung, M., Ellis, P., Loo, P.V., McLaren, S., Wedge, D.C., Fullam, A., Alexandrov, L.B., Tubio, J.M., et al. (2015). High burden and pervasive positive selection of somatic mutations in normal human skin. *Science* 348, 880–886.

- [360] Nielsen, R., Paul, J.S., Albrechtsen, A., and Song, Y.S. (2011). Genotype and SNP calling from next-generation sequencing data. *Nat. Rev. Genet.* 12, 443–451.
- [361] McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytzky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., et al. (2010). The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 20, 1297–1303.
- [362] DePristo, M.A., Banks, E., Poplin, R., Garimella, K.V., Maguire, J.R., Hartl, C., Philippakis, A.A., del Angel, G., Rivas, M.A., Hanna, M., et al. (2011). A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet.* 43, 491–498.
- [363] Liu, X., Han, S., Wang, Z., Gelernter, J., and Yang, B.-Z. (2013a). Variant callers for next-generation sequencing data: a comparison study. *PloS One* 8, e75619.
- [364] Pirooznia, M., Kramer, M., Parla, J., Goes, F.S., Potash, J.B., McCombie, W.R., and Zandi, P.P. (2014). Validation and assessment of variant calling pipelines for next-generation sequencing. *Hum. Genomics* 8, 14.
- [365] Yi, M., Zhao, Y., Jia, L., He, M., Kebebew, E., and Stephens, R.M. (2014). Performance comparison of SNP detection tools with illumina exome sequencing data--an assessment using both family pedigree information and sample-matched SNP array data. *Nucleic Acids Res.* 42, e101.
- [366] Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and 1000 Genome Project Data Processing Subgroup (2009). The Sequence Alignment/Map format and SAMtools. *Bioinforma. Oxf. Engl.* 25, 2078–2079.
- [367] Li, H. (2011). A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinforma. Oxf. Engl.* 27, 2987–2993.
- [368] Lam, H.Y.K., Clark, M.J., Chen, R., Chen, R., Natsoulis, G., O’Huallachain, M., Dewey, F.E., Habegger, L., Ashley, E.A., Gerstein, M.B., et al. (2012a). Performance comparison of whole-genome sequencing platforms. *Nat. Biotechnol.* 30, 78–82.
- [369] O’Rawe, J., Jiang, T., Sun, G., Wu, Y., Wang, W., Hu, J., Bodily, P., Tian, L., Hakonarson, H., Johnson, W.E., et al. (2013). Low concordance of multiple variant-calling pipelines: practical implications for exome and genome sequencing. *Genome Med.* 5, 28.
- [370] Albers, C.A., Lunter, G., MacArthur, D.G., McVean, G., Ouwehand, W.H., and Durbin, R. (2011). Dindel: accurate indel calls from short-read data. *Genome Res.* 21, 961–973.
- [371] Emde, A.-K., Schulz, M.H., Weese, D., Sun, R., Vingron, M., Kalscheuer, V.M., Haas, S.A., and Reinert, K. (2012). Detecting genomic indel variants with exact breakpoints in single- and paired-end sequencing data using SplazerS. *Bioinforma. Oxf. Engl.* 28, 619–627.
- [372] Kozarewa, I., Ning, Z., Quail, M.A., Sanders, M.J., Berriman, M., and Turner, D.J. (2009). Amplification-free Illumina sequencing-library preparation facilitates improved mapping and assembly of (G+C)-biased genomes. *Nat. Methods* 6, 291–295.

- [373] Mader, R.M., Schmidt, W.M., Sedivy, R., Rizovski, B., Braun, J., Kalipciyan, M., Exner, M., Steger, G.G., and Mueller, M.W. (2001). Reverse transcriptase template switching during reverse transcriptase-polymerase chain reaction: artificial generation of deletions in ribonucleotide reductase mRNA. *J. Lab. Clin. Med.* 137, 422–428.
- [374] Cocquet, J., Chong, A., Zhang, G., and Veitia, R.A. (2006). Reverse transcriptase template switching and false alternative transcripts. *Genomics* 88, 127–131.
- [375] Roy, S.W., and Irimia, M. (2008). When good transcripts go bad: artifactual RT-PCR “splicing” and genome analysis. *BioEssays News Rev. Mol. Cell. Dev. Biol.* 30, 601–605.
- [376] Armour, C.D., Castle, J.C., Chen, R., Babak, T., Loerch, P., Jackson, S., Shah, J.K., Dey, J., Rohl, C.A., Johnson, J.M., et al. (2009). Digital transcriptome profiling using selective hexamer priming for cDNA synthesis. *Nat. Methods* 6, 647–649.
- [377] Hansen, K.D., Brenner, S.E., and Dudoit, S. (2010). Biases in Illumina transcriptome sequencing caused by random hexamer priming. *Nucleic Acids Res.* 38, e131.
- [378] Dohm, J.C., Lottaz, C., Borodina, T., and Himmelbauer, H. (2008). Substantial biases in ultra-short read data sets from high-throughput DNA sequencing. *Nucleic Acids Res.* 36, e105.
- [379] Minoche, A.E., Dohm, J.C., and Himmelbauer, H. (2011). Evaluation of genomic high-throughput sequencing data generated on Illumina HiSeq and genome analyzer systems. *Genome Biol.* 12, R112.
- [380] Seyednasrollah, F., Laiho, A., and Elo, L.L. (2015). Comparison of software packages for detecting differential expression in RNA-seq studies. *Brief. Bioinform.* 16, 59–70.
- [381] Trapnell, C., Williams, B.A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M.J., Salzberg, S.L., Wold, B.J., and Pachter, L. (2010). Transcript assembly and abundance estimation from RNA-Seq reveals thousands of new transcripts and switching among isoforms. *Nat. Biotechnol.* 28, 511–515.
- [382] Trapnell, C., Hendrickson, D.G., Sauvageau, M., Goff, L., Rinn, J.L., and Pachter, L. (2013). Differential analysis of gene regulation at transcript resolution with RNA-seq. *Nat. Biotechnol.* 31, 46–53.
- [383] Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D.R., Pimentel, H., Salzberg, S.L., Rinn, J.L., and Pachter, L. (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat. Protoc.* 7, 562–578.
- [384] Nookaew, I., Papini, M., Pornputtpong, N., Scalcinati, G., Fagerberg, L., Uhlén, M., and Nielsen, J. (2012). A comprehensive comparison of RNA-Seq-based transcriptome analysis from reads to differential gene expression and cross-comparison with microarrays: a case study in *Saccharomyces cerevisiae*. *Nucleic Acids Res.* 40, 10084–10097.
- [385] Zhang, Z.H., Jhaveri, D.J., Marshall, V.M., Bauer, D.C., Edson, J., Narayanan, R.K., Robinson, G.J., Lundberg, A.E., Bartlett, P.F., Wray, N.R., et al. (2014). A Comparative Study of Techniques for Differential Expression Analysis on RNA-Seq Data. *PLoS ONE* 9, e103207.

- [386] Chen, H.-W., Yu, S.-L., Chen, W.-J., Yang, P.-C., Chien, C.-T., Chou, H.-Y., Li, H.-N., Peck, K., Huang, C.-H., Lin, F.-Y., et al. (2004). Dynamic changes of gene expression profiles during postnatal development of the heart in mice. *Heart Br. Card. Soc.* 90, 927–934.
- [387] Liu, Y., Zhou, J., and White, K.P. (2013b). RNA-seq differential expression studies: more sequence, or more replication? *Bioinformatics* btt688.
- [388] Grunau, C., Clark, S.J., and Rosenthal, A. (2001). Bisulfite genomic sequencing: systematic investigation of critical experimental parameters. *Nucleic Acids Res.* 29, E65–65.
- [389] Genereux, D.P., Johnson, W.C., Burden, A.F., Stöger, R., and Laird, C.D. (2008). Errors in the bisulfite conversion of DNA: modulating inappropriate- and failed-conversion frequencies. *Nucleic Acids Res.* 36, e150.
- [390] Holmes, E.E., Jung, M., Meller, S., Leisse, A., Sailer, V., Zech, J., Mengdehl, M., Garbe, L.-A., Uhl, B., Kristiansen, G., et al. (2014). Performance Evaluation of Kits for Bisulfite-Conversion of DNA from Tissues, Cell Lines, FFPE Tissues, Aspirates, Lavages, Effusions, Plasma, Serum, and Urine. *PLoS ONE* 9, e93933.
- [391] Raizis, A.M., Schmitt, F., and Jost, J.P. (1995). A bisulfite method of 5-methylcytosine mapping that minimizes template degradation. *Anal. Biochem.* 226, 161–166.
- [392] Tanaka, K., and Okamoto, A. (2007). Degradation of DNA by bisulfite treatment. *Bioorg. Med. Chem. Lett.* 17, 1912–1915.
- [393] Shiraishi, M., and Hayatsu, H. (2004). High-speed conversion of cytosine to uracil in bisulfite genomic sequencing analysis of DNA methylation. *DNA Res. Int. J. Rapid Publ. Rep. Genes Genomes* 11, 409–415.
- [394] Sandoval, J., Heyn, H., Moran, S., Serra-Musach, J., Pujana, M.A., Bibikova, M., and Esteller, M. (2011). Validation of a DNA methylation microarray for 450,000 CpG sites in the human genome. *Epigenetics* 6, 692–702.
- [395] Stirzaker, C., Taberlay, P.C., Statham, A.L., and Clark, S.J. (2014). Mining cancer methylomes: prospects and challenges. *Trends Genet.* 30, 75–84.
- [396] Dedeurwaerder, S., Defrance, M., Bizet, M., Calonne, E., Bontempi, G., and Fuks, F. (2013). A comprehensive overview of Infinium HumanMethylation450 data processing. *Brief. Bioinform.* bbt054.
- [397] Price, M.E., Cotton, A.M., Lam, L.L., Farré, P., Emberly, E., Brown, C.J., Robinson, W.P., and Kobor, M.S. (2013). Additional annotation enhances potential for biologically-relevant analysis of the Illumina Infinium HumanMethylation450 BeadChip array. *Epigenetics Chromatin* 6, 4.
- [398] Zhang, X., Mu, W., and Zhang, W. (2012). On the Analysis of the Illumina 450k Array Data: Probes Ambiguously Mapped to the Human Genome. *Front. Genet.* 3.
- [399] Leek, J.T., Scharpf, R.B., Bravo, H.C., Simcha, D., Langmead, B., Johnson, W.E., Geman, D., Baggerly, K., and Irizarry, R.A. (2010). Tackling the widespread and critical impact of batch effects in high-throughput data. *Nat. Rev. Genet.* 11.

- [400] Morris, T.J., and Beck, S. (2015). Analysis pipelines and packages for Infinium HumanMethylation450 BeadChip (450k) data. *Methods San Diego Calif* 72, 3–8.
- [401] Eijk, K.R. van, Jong, S. de, Boks, M.P., Langeveld, T., Colas, F., Veldink, J.H., Kovel, C.G. de, Janson, E., Strengman, E., Langfelder, P., et al. (2012). Genetic analysis of DNA methylation and gene expression levels in whole blood of healthy human subjects. *BMC Genomics* 13, 636.
- [402] Dedeurwaerder, S., Defrance, M., Calonne, E., Denis, H., Sotiriou, C., and Fuks, F. (2011). Evaluation of the Infinium Methylation 450K technology. *Epigenomics* 3, 771–784.
- [403] Patrushev, L.I., and Kovalenko, T.F. (2014). Functions of noncoding sequences in mammalian genomes. *Biochem. Biokhimiia* 79, 1442–1469.
- [404] Roessler, J., Ammerpohl, O., Gutwein, J., Hasemeier, B., Anwar, S.L., Kreipe, H., and Lehmann, U. (2012). Quantitative cross-validation and content analysis of the 450k DNA methylation array from Illumina, Inc. *BMC Res. Notes* 5, 210.
- [405] Izzi, B., Decallonne, B., Devriendt, K., Bouillon, R., Vanderschueren, D., Levtschenko, E., de Zegher, F., Van den Bruel, A., Lambrechts, D., Van Geet, C., et al. (2010). A new approach to imprinting mutation detection in GNAS by Sequenom EpiTYPER system. *Clin. Chim. Acta Int. J. Clin. Chem.* 411, 2033–2039.
- [406] Ehrich, M., Field, J.K., Liloglou, T., Xinarianos, G., Oeth, P., Nelson, M.R., Cantor, C.R., and van den Boom, D. (2006). Cytosine methylation profiles as a molecular marker in non-small cell lung cancer. *Cancer Res.* 66, 10911–10918.
- [407] Nygren, A.O.H., Dean, J., Jensen, T.J., Kruse, S., Kwong, W., van den Boom, D., and Ehrich, M. (2010). Quantification of fetal DNA by use of methylation-based DNA discrimination. *Clin. Chem.* 56, 1627–1635.
- [408] Eskedal, L., Hagemo, P., Eskild, A., Aamodt, G., Seiler, K.S., and Thaulow, E. (2004). A population-based study of extra-cardiac anomalies in children with congenital cardiac malformations. *Cardiol. Young* 14, 600–607.
- [409] Meberg, A., Hals, J., and Thaulow, E. (2007). Congenital heart defects—chromosomal anomalies, syndromes and extracardiac malformations. *Acta Pædiatrica* 96, 1142–1145.
- [410] Lesk, A.M. (1995). NAD-binding domains of dehydrogenases. *Curr. Opin. Struct. Biol.* 5, 775–783.
- [411] Moriarity, J.L., Hurt, K.J., Resnick, A.C., Storm, P.B., Laroy, W., Schnaar, R.L., and Snyder, S.H. (2002). UDP-glucuronate decarboxylase, a key enzyme in proteoglycan synthesis: cloning, characterization, and localization. *J. Biol. Chem.* 277, 16968–16975.
- [412] Hubmacher, D., and Apte, S.S. (2013). The biology of the extracellular matrix: novel insights. *Curr. Opin. Rheumatol.* 25, 65–70.
- [413] Lockhart, M., Wirrig, E., Phelps, A., and Wessels, A. (2011). Extracellular Matrix and Heart Development. *Birt. Defects Res. A. Clin. Mol. Teratol.* 91, 535–550.
- [414] Watanabe, H., and Yamada, Y. (1999). Mice lacking link protein develop dwarfism and craniofacial abnormalities. *Nat. Genet.* 21, 225–229.

- [415] Wirrig, E.E., Snarr, B.S., Chintalapudi, M.R., O'neal, J.L., Phelps, A.L., Barth, J.L., Fresco, V.M., Kern, C.B., Mjaatvedt, C.H., Toole, B.P., et al. (2007). Cartilage link protein 1 (Crtl1), an extracellular matrix component playing an important role in heart development. *Dev. Biol.* 310, 291–303.
- [416] Ng, A., Wong, M., Viviano, B., Erlich, J.M., Alba, G., Pflederer, C., Jay, P.Y., and Saunders, S. (2009a). Loss of glypican-3 function causes growth factor-dependent defects in cardiac and coronary vascular development. *Dev. Biol.* 335, 208–215.
- [417] Carmel, I., Tal, S., Vig, I., and Ast, G. (2004). Comparative analysis detects dependencies among the 5' splice-site positions. *RNA N. Y. N* 10, 828–840.
- [418] Pereira, F.A., Qiu, Y., Tsai, M.J., and Tsai, S.Y. (1995). Chicken ovalbumin upstream promoter transcription factor (COUP-TF): expression during mouse embryogenesis. *J. Steroid Biochem. Mol. Biol.* 53, 503–508.
- [419] Tsai, S.Y., and Tsai, M.J. (1997). Chick ovalbumin upstream promoter-transcription factors (COUP-TFs): coming of age. *Endocr. Rev.* 18, 229–240.
- [420] Pereira, F.A., Qiu, Y., Zhou, G., Tsai, M.-J., and Tsai, S.Y. (1999). The orphan nuclear receptor COUP-TFII is required for angiogenesis and heart development. *Genes Dev.* 13, 1037–1049.
- [421] Kovacic, J.C., Mercader, N., Torres, M., Boehm, M., and Fuster, V. (2012). Epithelial-to-Mesenchymal and Endothelial-to-Mesenchymal Transition From Cardiovascular Development to Disease. *Circulation* 125, 1795–1808.
- [422] Markwald, R.R., Fitzharris, T.P., and Manasek, F.J. (1977). Structural development of endocardial cushions. *Am. J. Anat.* 148, 85–119.
- [423] Kisanuki, Y.Y., Hammer, R.E., Miyazaki, J., Williams, S.C., Richardson, J.A., and Yanagisawa, M. (2001). Tie2-Cre transgenic mice: a new model for endothelial cell-lineage analysis in vivo. *Dev. Biol.* 230, 230–242.
- [424] von Gise, A., and Pu, W.T. (2012). Endocardial and epicardial epithelial to mesenchymal transitions in heart development and disease. *Circ. Res.* 110, 1628–1645.
- [425] Lin, F.-J., You, L.-R., Yu, C.-T., Hsu, W.-H., Tsai, M.-J., and Tsai, S.Y. (2012). Endocardial cushion morphogenesis and coronary vessel development require chicken ovalbumin upstream promoter-transcription factor II. *Arterioscler. Thromb. Vasc. Biol.* 32, e135–e146.
- [426] Basson, C.T., Cowley, G.S., Solomon, S.D., Weissman, B., Poznanski, A.K., Traill, T.A., Seidman, J.G., and Seidman, C.E. (1994). The clinical and genetic spectrum of the Holt-Oram syndrome (heart-hand syndrome). *N. Engl. J. Med.* 330, 885–891.
- [427] Zhu, H., Kartiko, S., and Finnell, R.H. (2009). Importance of gene-environment interactions in the etiology of selected birth defects. *Clin. Genet.* 75, 409–423.
- [428] Trapani, J.A. (1995). Target cell apoptosis induced by cytotoxic T cells and natural killer cells involves synergy between the pore-forming protein, perforin, and the serine protease, granzyme B. *Aust. N. Z. J. Med.* 25, 793–799.

- [429] Voskoboinik, I., Dunstone, M.A., Baran, K., Whisstock, J.C., and Trapani, J.A. (2010). Perforin: structure, function, and role in human immunopathology. *Immunol. Rev.* 235, 35–54.
- [430] Felzen, B., Berke, G., Rosen, D., Coleman, R., Tschopp, J., Young, J.D., and Binah, O. (1994). Effects of purified perforin and granzyme A from cytotoxic T lymphocytes on guinea pig ventricular myocytes. *Cardiovasc. Res.* 28, 643–649.
- [431] Sabbah, H.N. (2000). Apoptotic cell death in heart failure. *Cardiovasc. Res.* 45, 704–712.
- [432] Gill, C., Mestral, R., and Samali, A. (2002). Losing heart: the role of apoptosis in heart disease—a novel therapeutic target? *FASEB J.* 16, 135–146.
- [433] Whelan, R.S., Kaplinskiy, V., and Kitsis, R.N. (2010). Cell death in the pathogenesis of heart disease: mechanisms and significance. *Annu. Rev. Physiol.* 72, 19–44.
- [434] Chiong, M., Wang, Z.V., Pedrozo, Z., Cao, D.J., Troncoso, R., Ibacache, M., Criollo, A., Nemchenko, A., Hill, J.A., and Lavandero, S. (2011). Cardiomyocyte death: mechanisms and translational implications. *Cell Death Dis.* 2, e244.
- [435] Konstantinidis, K., Whelan, R.S., and Kitsis, R.N. (2012). Mechanisms of Cell Death in Heart Disease. *Arterioscler. Thromb. Vasc. Biol.* 32, 1552–1562.
- [436] Shur, N., Cowan, J., and Wheeler, P.G. (2003). Craniosynostosis and congenital heart anomalies associated with a maternal deletion of 15q15-22.1. *Am. J. Med. Genet. A.* 120A, 542–546.
- [437] Chang, P., Coughlin, M., and Mitchison, T.J. (2005). Tankyrase-1 polymerization of poly(ADP-ribose) is required for spindle structure and function. *Nat. Cell Biol.* 7, 1133–1139.
- [438] Dynek, J.N., and Smith, S. (2004). Resolution of sister telomere association is required for progression through mitosis. *Science* 304, 97–100.
- [439] Bowles, K.R., Gajarski, R., Porter, P., Goytia, V., Bachinski, L., Roberts, R., Pignatelli, R., and Towbin, J.A. (1996). Gene mapping of familial autosomal dominant dilated cardiomyopathy to chromosome 10q21-23. *J. Clin. Invest.* 98, 1355–1360.
- [440] Spielmann, M., and Klopocki, E. (2013). CNVs of noncoding cis-regulatory elements in human disease. *Curr. Opin. Genet. Dev.* 23, 249–256.
- [441] Lupski, J.R., and Stankiewicz, P. (2005). Genomic disorders: molecular mechanisms for rearrangements and conveyed phenotypes. *PLoS Genet.* 1, e49.
- [442] Lander, J., and Ware, S.M. (2014). Copy Number Variation in Congenital Heart Defects. *Curr. Genet. Med. Rep.* 2, 168–178.
- [443] McFadden, D.G., Barbosa, A.C., Richardson, J.A., Schneider, M.D., Srivastava, D., and Olson, E.N. (2005). The Hand1 and Hand2 transcription factors regulate expansion of the embryonic cardiac ventricles in a gene dosage-dependent manner. *Dev. Camb. Engl.* 132, 189–201.
- [444] Washington Smoak, I., Byrd, N.A., Abu-Issa, R., Goddeeris, M.M., Anderson, R., Morris, J., Yamamura, K., Klingensmith, J., and Meyers, E.N. (2005). Sonic hedgehog is required for cardiac outflow tract and neural crest cell development. *Dev. Biol.* 283, 357–372.

- [445] Goddeeris, M.M., Rho, S., Petiet, A., Davenport, C.L., Johnson, G.A., Meyers, E.N., and Klingensmith, J. (2008). Intracardiac septation requires hedgehog-dependent cellular contributions from outside the heart. *Development* 135, 1887–1895.
- [446] Thomas, N.A., Koudijs, M., van Eeden, F.J.M., Joyner, A.L., and Yelon, D. (2008). Hedgehog signaling plays a cell-autonomous role in maximizing cardiac developmental potential. *Dev. Camb. Engl.* 135, 3789–3799.
- [447] Wong, K.S., Rehn, K., Palencia-Desai, S., Kohli, V., Hunter, W., Uhl, J.D., Rost, M.S., and Sumanas, S. (2012). Hedgehog signaling is required for differentiation of endocardial progenitors in zebrafish. *Dev. Biol.* 361, 377–391.
- [448] Quijada, L., Callejo, A., Torroja, C., and Guerrero, I. (2000). The Patched Receptor: Switching On/Off the Hedgehog Signaling Pathway.
- [449] Chiang, C., Litingtung, Y., Lee, E., Young, K.E., Corden, J.L., Westphal, H., and Beachy, P.A. (1996). Cyclopia and defective axial patterning in mice lacking Sonic hedgehog gene function. *Nature* 383, 407–413.
- [450] Tsukui, T., Capdevila, J., Tamura, K., Ruiz-Lozano, P., Rodriguez-Esteban, C., Yonei-Tamura, S., Magallón, J., Chandraratna, R.A., Chien, K., Blumberg, B., et al. (1999). Multiple left-right asymmetry defects in *Shh*(^{-/-}) mutant mice unveil a convergence of the *shh* and retinoic acid pathways in the control of *Lefty-1*. *Proc. Natl. Acad. Sci. U. S. A.* 96, 11376–11381.
- [451] High, F.A., and Epstein, J.A. (2008). The multifaceted role of Notch in cardiac development and disease. *Nat. Rev. Genet.* 9, 49–61.
- [452] de la Pompa, J.L., and Epstein, J.A. (2012). Coordinating tissue interactions: Notch signaling in cardiac development and disease. *Dev. Cell* 22, 244–254.
- [453] Krebs, L.T., Xue, Y., Norton, C.R., Shutter, J.R., Maguire, M., Sundberg, J.P., Gallahan, D., Closson, V., Kitajewski, J., Callahan, R., et al. (2000). Notch signaling is essential for vascular morphogenesis in mice. *Genes Dev.* 14, 1343–1352.
- [454] Grieskamp, T., Rudat, C., Lütke, T.H.-W., Norden, J., and Kispert, A. (2011). Notch signaling regulates smooth muscle differentiation of epicardium-derived cells. *Circ. Res.* 108, 813–823.
- [455] Chang, A.C.Y., Fu, Y., Garside, V.C., Niessen, K., Chang, L., Fuller, M., Setiadi, A., Smrz, J., Kyle, A., Minchinton, A., et al. (2011). Notch initiates the endothelial-to-mesenchymal transition in the atrioventricular canal through autocrine activation of soluble guanylyl cyclase. *Dev. Cell* 21, 288–300.
- [456] Riley, M.F., McBride, K.L., and Cole, S.E. (2011). NOTCH1 missense alleles associated with left ventricular outflow tract defects exhibit impaired receptor processing and defective EMT. *Biochim. Biophys. Acta* 1812, 121–129.
- [457] Liu, J., Dong, F., Jeong, J., Masuda, T., and Lobe, C.G. (2014). Constitutively active Notch1 signaling promotes endothelial-mesenchymal transition in a conditional transgenic mouse model. *Int. J. Mol. Med.* 34, 669–676.

- [458] Garg, V., Muth, A.N., Ransom, J.F., Schluterman, M.K., Barnes, R., King, I.N., Grossfeld, P.D., and Srivastava, D. (2005). Mutations in NOTCH1 cause aortic valve disease. *Nature* 437, 270–274.
- [459] Mohamed, S.A., Aherrahrou, Z., Liptau, H., Erasmi, A.W., Hagemann, C., Wrobel, S., Borzym, K., Schunkert, H., Sievers, H.H., and Erdmann, J. (2006). Novel missense mutations (p.T596M and p.P1797H) in NOTCH1 in patients with bicuspid aortic valve. *Biochem. Biophys. Res. Commun.* 345, 1460–1465.
- [460] McKellar, S.H., Tester, D.J., Yagubyan, M., Majumdar, R., Ackerman, M.J., and Sundt, T.M. (2007). Novel NOTCH1 mutations in patients with bicuspid aortic valve disease and thoracic aortic aneurysms. *J. Thorac. Cardiovasc. Surg.* 134, 290–296.
- [461] Foffa, I., Ait Ali, L., Panesi, P., Mariani, M., Festa, P., Botto, N., Vecoli, C., and Andreassi, M.G. (2013). Sequencing of NOTCH1, GATA5, TGFBR1 and TGFBR2 genes in familial cases of bicuspid aortic valve. *BMC Med. Genet.* 14, 44.
- [462] Zhang, A., Yeung, P.L., Li, C.-W., Tsai, S.-C., Dinh, G.K., Wu, X., Li, H., and Chen, J.D. (2004). Identification of a Novel Family of Ankyrin Repeats Containing Cofactors for p160 Nuclear Receptor Coactivators. *J. Biol. Chem.* 279, 33799–33805.
- [463] Brancati, F., D’Avanzo, M.G., Digilio, M.C., Sarkozy, A., Biondi, M., De Brasi, D., Mingarelli, R., and Dallapiccola, B. (2004). KBG syndrome in a cohort of Italian patients. *Am. J. Med. Genet. A.* 131, 144–149.
- [464] Lemerrier, C., To, R.Q., Carrasco, R.A., and Konieczny, S.F. (1998). The basic helix-loop-helix transcription factor Mist1 functions as a transcriptional repressor of myoD. *EMBO J.* 17, 1412–1422.
- [465] Vecoli, C., Pulignani, S., Foffa, I., and Andreassi, M.G. (2014). Congenital Heart Disease: The Crossroads of Genetics, Epigenetics and Environment. *Curr. Genomics* 15, 390–399.
- [466] Rajasingh, J. (2012). Reprogramming of somatic cells. *Prog. Mol. Biol. Transl. Sci.* 111, 51–82.
- [467] Irizarry, R.A., Ladd-Acosta, C., Wen, B., Wu, Z., Montano, C., Onyango, P., Cui, H., Gabo, K., Rongione, M., Webster, M., et al. (2009). The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. *Nat. Genet.* 41, 178–186.
- [468] Cavodeassi, F., Modolell, J., and Gómez-Skarmeta, J.L. (2001). The Iroquois family of genes: from body building to neural patterning. *Dev. Camb. Engl.* 128, 2847–2855.
- [469] Bao, Z.Z., Bruneau, B.G., Seidman, J.G., Seidman, C.E., and Cepko, C.L. (1999). Regulation of chamber-specific gene expression in the developing heart by Irx4. *Science* 283, 1161–1164.
- [470] Bruneau, B.G., Bao, Z.Z., Tanaka, M., Schott, J.J., Izumo, S., Cepko, C.L., Seidman, J.G., and Seidman, C.E. (2000). Cardiac expression of the ventricle-specific homeobox gene Irx4 is modulated by Nkx2-5 and dHand. *Dev. Biol.* 217, 266–277.
- [471] R J Garriock, S.A.V. (2001). Developmental expression of the Xenopus Iroquois-family homeobox genes, Irx4 and Irx5. *Dev. Genes Evol.* 211, 257–260.

- [472] Bruneau, B.G., Bao, Z.Z., Fatkin, D., Xavier-Neto, J., Georgakopoulos, D., Maguire, C.T., Berul, C.I., Kass, D.A., Kuroski-de Bold, M.L., de Bold, A.J., et al. (2001b). Cardiomyopathy in *Irx4*-deficient mice is preceded by abnormal ventricular gene expression. *Mol. Cell. Biol.* 21, 1730–1736.
- [473] Illingworth, R.S., Gruenewald-Schneider, U., Webb, S., Kerr, A.R.W., James, K.D., Turner, D.J., Smith, C., Harrison, D.J., Andrews, R., and Bird, A.P. (2010). Orphan CpG islands identify numerous conserved promoters in the mammalian genome. *PLoS Genet.* 6, e1001134.
- [474] Maunakea, A.K., Nagarajan, R.P., Bilenky, M., Ballinger, T.J., D'Souza, C., Fouse, S.D., Johnson, B.E., Hong, C., Nielsen, C., Zhao, Y., et al. (2010). Conserved role of intragenic DNA methylation in regulating alternative promoters. *Nature* 466, 253–257.
- [475] Beiman, M., Shilo, B.Z., and Volk, T. (1996). Heartless, a *Drosophila* FGF receptor homolog, is essential for cell migration and establishment of several mesodermal lineages. *Genes Dev.* 10, 2993–3002.
- [476] Gryzik, T., and Müller, H.-A.J. (2004). FGF8-like1 and FGF8-like2 encode putative ligands of the FGF receptor Htl and are required for mesoderm migration in the *Drosophila* gastrula. *Curr. Biol. CB* 14, 659–667.
- [477] Stathopoulos, A., Tam, B., Ronshaugen, M., Frasch, M., and Levine, M. (2004). *pyramus* and *thisbe*: FGF genes that pattern the mesoderm of *Drosophila* embryos. *Genes Dev.* 18, 687–699.
- [478] Reifers, F., Walsh, E.C., Léger, S., Stainier, D.Y., and Brand, M. (2000). Induction and differentiation of the zebrafish heart requires fibroblast growth factor 8 (*fgf8/acerebellar*). *Dev. Camb. Engl.* 127, 225–235.
- [479] Alsan, B.H., and Schultheiss, T.M. (2002). Regulation of avian cardiogenesis by *Fgf8* signaling. *Dev. Camb. Engl.* 129, 1935–1943.
- [480] Abu-Issa, R., Smyth, G., Smoak, I., Yamamura, K., and Meyers, E.N. (2002). *Fgf8* is required for pharyngeal arch and cardiovascular development in the mouse. *Dev. Camb. Engl.* 129, 4613–4625.
- [481] Frank, D.U., Fotheringham, L.K., Brewer, J.A., Muglia, L.J., Tristani-Firouzi, M., Capecchi, M.R., and Moon, A.M. (2002). An *Fgf8* mouse mutant phenocopies human 22q11 deletion syndrome. *Dev. Camb. Engl.* 129, 4591–4603.
- [482] Marguerie, A., Bajolle, F., Zaffran, S., Brown, N.A., Dickson, C., Buckingham, M.E., and Kelly, R.G. (2006). Congenital heart defects in *Fgfr2-IIIb* and *Fgf10* mutant mice. *Cardiovasc. Res.* 71, 50–60.
- [483] Wiench, M., John, S., Baek, S., Johnson, T.A., Sung, M.-H., Escobar, T., Simmons, C.A., Pearce, K.H., Biddie, S.C., Sabo, P.J., et al. (2011). DNA methylation status predicts cell type-specific enhancer activity. *EMBO J.* 30, 3028–3039.
- [484] Ng, S.Y., Wong, C.K., and Tsang, S.Y. (2010). Differential gene expressions in atrial and ventricular myocytes: insights into the road of applying embryonic stem cell-derived cardiomyocytes for future therapies. *Am. J. Physiol. - Cell Physiol.* 299, C1234–C1249.

- [485] He, W., Jia, Y., and Takimoto, K. (2009). Interaction between transcription factors Iroquois proteins 4 and 5 controls cardiac potassium channel Kv4.2 gene transcription. *Cardiovasc. Res.* 81, 64–71.
- [486] Xu, H., Yi, B.A., Wu, H., Bock, C., Gu, H., Lui, K.O., Park, J.-H.C., Shao, Y., Riley, A.K., Domian, I.J., et al. (2012). Highly efficient derivation of ventricular cardiomyocytes from induced pluripotent stem cells with a distinct epigenetic signature. *Cell Res.* 22, 142–154.
- [487] Riggs, A.D. (1975). X inactivation, differentiation, and DNA methylation. *Cytogenet. Cell Genet.* 14, 9–25.
- [488] Hsieh, C.L. (1994). Dependence of transcriptional repression on CpG methylation density. *Mol. Cell. Biol.* 14, 5487–5494.
- [489] Walsh, C.P., and Bestor, T.H. (1999). Cytosine methylation and mammalian development. *Genes Dev.* 13, 26–34.
- [490] Shen, L., Kondo, Y., Guo, Y., Zhang, J., Zhang, L., Ahmed, S., Shu, J., Chen, X., Waterland, R.A., and Issa, J.-P.J. (2007). Genome-Wide Profiling of DNA Methylation Reveals a Class of Normally Methylated CpG Island Promoters. *PLoS Genet* 3, e181.
- [491] Mukhopadhyay, R., Yu, W., Whitehead, J., Xu, J., Lezcano, M., Pack, S., Kanduri, C., Kanduri, M., Ginja, V., Vostrov, A., et al. (2004). The Binding Sites for the Chromatin Insulator Protein CTCF Map to DNA Methylation-Free Domains Genome-Wide. *Genome Res.* 14, 1594–1602.
- [492] Lister, R., Pelizzola, M., Dowen, R.H., Hawkins, R.D., Hon, G., Tonti-Filippini, J., Nery, J.R., Lee, L., Ye, Z., Ngo, Q.-M., et al. (2009). Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* 462, 315–322.
- [493] Thurman, R.E., Rynes, E., Humbert, R., Vierstra, J., Maurano, M.T., Haugen, E., Sheffield, N.C., Stergachis, A.B., Wang, H., Vernot, B., et al. (2012). The accessible chromatin landscape of the human genome. *Nature* 489, 75–82.
- [494] Straussman, R., Nejman, D., Roberts, D., Steinfeld, I., Blum, B., Benvenisty, N., Simon, I., Yakhini, Z., and Cedar, H. (2009). Developmental programming of CpG island methylation profiles in the human genome. *Nat. Struct. Mol. Biol.* 16, 564–571.
- [495] Gebhard, C., Benner, C., Ehrich, M., Schwarzfischer, L., Schilling, E., Klug, M., Dietmaier, W., Thiede, C., Holler, E., Andreesen, R., et al. (2010). General Transcription Factor Binding at CpG Islands in Normal Cells Correlates with Resistance to De novo DNA Methylation in Cancer Cells. *Cancer Res.* 70, 1398–1407.
- [496] Seale, P., Bjork, B., Yang, W., Kajimura, S., Chin, S., Kuang, S., Scimè, A., Devarakonda, S., Conroe, H.M., Erdjument-Bromage, H., et al. (2008). PRDM16 controls a brown fat/skeletal muscle switch. *Nature* 454, 961–967.
- [497] Kajimura, S., Seale, P., Kubota, K., Lunsford, E., Frangioni, J.V., Gygi, S.P., and Spiegelman, B.M. (2009). Initiation of myoblast to brown fat switch by a PRDM16–C/EBP- β transcriptional complex. *Nature* 460, 1154–1158.

- [498] Selker, E.U., and Stevens, J.N. (1985). DNA methylation at asymmetric sites is associated with numerous transition mutations. *Proc. Natl. Acad. Sci. U. S. A.* 82, 8114–8118.
- [499] Cooper, D.N., and Youssoufian, H. (1988). The CpG dinucleotide and human genetic disease. *Hum. Genet.* 78, 151–155.
- [500] Rideout, W.M., Coetzee, G.A., Olumi, A.F., and Jones, P.A. (1990). 5-Methylcytosine as an endogenous mutagen in the human LDL receptor and p53 genes. *Science* 249, 1288–1290.
- [501] Meurs, K.M., and Kuan, M. (2011). Differential methylation of CpG sites in two isoforms of myosin binding protein C, an important hypertrophic cardiomyopathy gene. *Environ. Mol. Mutagen.* 52, 161–164.
- [502] Yang, X., Ewald, E.R., Huo, Y., Tamashiro, K.L., Salvatori, R., Sawa, A., Wand, G.S., and Lee, R.S. (2012). Glucocorticoid-induced loss of DNA methylation in non-neuronal cells and potential involvement of DNMT1 in epigenetic regulation of Fkbp5. *Biochem. Biophys. Res. Commun.* 420, 570–575.
- [503] Rönn, T., Poulsen, P., Hansson, O., Holmkvist, J., Almgren, P., Nilsson, P., Tuomi, T., Isomaa, B., Groop, L., Vaag, A., et al. (2008). Age influences DNA methylation and gene expression of COX7A1 in human skeletal muscle. *Diabetologia* 51, 1159–1168.
- [504] Wong, C.C.Y., Caspi, A., Williams, B., Craig, I.W., Houts, R., Ambler, A., Moffitt, T.E., and Mill, J. (2010). A longitudinal study of epigenetic variation in twins. *Epigenetics* 5, 516–526.
- [505] Schneider, E., Pliushch, G., El Hajj, N., Galetzka, D., Puhl, A., Schorsch, M., Frauenknecht, K., Riepert, T., Tresch, A., Müller, A.M., et al. (2010). Spatial, temporal and interindividual epigenetic variation of functionally important DNA methylation patterns. *Nucleic Acids Res.* 38, 3880–3890.
- [506] Fryer, A.A., Emes, R.D., Ismail, K.M.K., Haworth, K.E., Mein, C., Carroll, W.D., and Farrell, W.E. (2011). Quantitative, high-resolution epigenetic profiling of CpG loci identifies associations with cord blood plasma homocysteine and birth weight in humans. *Epigenetics* 6, 86–94.
- [507] Lam, L.L., Emberly, E., Fraser, H.B., Neumann, S.M., Chen, E., Miller, G.E., and Kobor, M.S. (2012b). Factors underlying variable DNA methylation in a human community cohort. *Proc. Natl. Acad. Sci.* 109, 7253.
- [508] Choi, S.H., Worswick, S., Byun, H.-M., Shear, T., Soussa, J.C., Wolff, E.M., Douer, D., Garcia-Manero, G., Liang, G., and Yang, A.S. (2009b). Changes in DNA methylation of tandem DNA repeats are different from interspersed repeats in cancer. *Int. J. Cancer J. Int. Cancer* 125, 723–729.
- [509] Feinberg, A.P., and Irizarry, R.A. (2010). Evolution in health and medicine Sackler colloquium: Stochastic epigenetic variation as a driving force of development, evolutionary adaptation, and disease. *Proc. Natl. Acad. Sci. U. S. A.* 107 Suppl 1, 1757–1764.
- [510] El-Maarri, O., Walier, M., Behne, F., van Üüm, J., Singer, H., Diaz-Lacava, A., Nüsgen, N., Niemann, B., Watzka, M., Reinsberg, J., et al. (2011). Methylation at global LINE-1 repeats in human blood are affected by gender but not by age or natural hormone cycles. *PLoS One* 6, e16252.

- [511] Coolen, M.W., Statham, A.L., Qu, W., Campbell, M.J., Henders, A.K., Montgomery, G.W., Martin, N.G., and Clark, S.J. (2011). Impact of the genome on the epigenome is manifested in DNA methylation patterns of imprinted regions in monozygotic and dizygotic twins. *PLoS One* 6, e25590.
- [512] Cloherty, J.P., Eichenwald, E.C., and Hansen, A.R. (2012). *Manual of Neonatal Care* (Lippincott Williams & Wilkins).
- [513] Rosenfeld, M.G., and Glass, C.K. (2001). Coregulator Codes of Transcriptional Regulation by Nuclear Receptors. *J. Biol. Chem.* 276, 36865–36868.
- [514] Serra-Juhé, C., Cuscó, I., Homs, A., Flores, R., Torán, N., and Pérez-Jurado, L.A. (2015). DNA methylation abnormalities in congenital heart disease. *Epigenetics* 10, 167–177.
- [515] Lokk, K., Modhukur, V., Rajashekar, B., Märten, K., Mägi, R., Kolde, R., Koltšina, M., Nilsson, T.K., Vilo, J., Salumets, A., et al. (2014). DNA methylome profiling of human tissues identifies global and tissue-specific methylation patterns. *Genome Biol.* 15, r54.
- [516] Grimsley, C.M., Kinchen, J.M., Tosello-Tramont, A.-C., Brugnera, E., Haney, L.B., Lu, M., Chen, Q., Klingele, D., Hengartner, M.O., and Ravichandran, K.S. (2004). Dock180 and ELMO1 proteins cooperate to promote evolutionarily conserved Rac-dependent cell migration. *J. Biol. Chem.* 279, 6087–6097.
- [517] Li, X., Gao, X., Liu, G., Xiong, W., Wu, J., and Rao, Y. (2008). Netrin signal transduction and the guanine nucleotide exchange factor DOCK180 in attractive signaling. *Nat. Neurosci.* 11, 28–35.
- [518] Laurin, M., Fradet, N., Blangy, A., Hall, A., Vuori, K., and Côté, J.-F. (2008). The atypical Rac activator Dock180 (Dock1) regulates myoblast fusion in vivo. *Proc. Natl. Acad. Sci. U. S. A.* 105, 15446–15451.
- [519] Sanematsu, F., Hirashima, M., Laurin, M., Takii, R., Nishikimi, A., Kitajima, K., Ding, G., Noda, M., Murata, Y., Tanaka, Y., et al. (2010). DOCK180 Is a Rac Activator That Regulates Cardiovascular Development by Acting Downstream of CXCR4. *Circ. Res.* 107, 1102–1105.
- [520] Schäker, K., Bartsch, S., Patry, C., Stoll, S.J., Hillebrands, J.-L., Wieland, T., and Kroll, J. (2015). The bipartite rac1 Guanine nucleotide exchange factor engulfment and cell motility 1/dedicator of cytokinesis 180 (elmo1/dock180) protects endothelial cells from apoptosis in blood vessel development. *J. Biol. Chem.* 290, 6408–6418.
- [521] Grigoryan, T., Wend, P., Klaus, A., and Birchmeier, W. (2008). Deciphering the function of canonical Wnt signals in development and disease: conditional loss- and gain-of-function mutations of β -catenin in mice. *Genes Dev.* 22, 2308–2341.
- [522] Huelsken, J., Vogel, R., Brinkmann, V., Erdmann, B., Birchmeier, C., and Birchmeier, W. (2000). Requirement for β -Catenin in Anterior-Posterior Axis Formation in Mice. *J. Cell Biol.* 148, 567–578.
- [523] Huh, S.-H., and Ornitz, D.M. (2010). β -catenin deficiency causes DiGeorge syndrome-like phenotypes through regulation of Tbx1. *Dev. Camb. Engl.* 137, 1137–1147.

- [524] Goulter, A.B., Goddard, M.J., Allen, J.C., and Clark, K.L. (2004). ACE2 gene expression is up-regulated in the human failing heart. *BMC Med.* 2, 19.
- [525] Boehm, M., and Nabel, E.G. (2002). Angiotensin-converting enzyme 2--a new cardiac regulator. *N. Engl. J. Med.* 347, 1795–1797.
- [526] Crackower, M.A., Sarao, R., Oudit, G.Y., Yagil, C., Kozieradzki, I., Scanga, S.E., Oliveira-dos-Santos, A.J., da Costa, J., Zhang, L., Pei, Y., et al. (2002). Angiotensin-converting enzyme 2 is an essential regulator of heart function. *Nature* 417, 822–828.
- [527] Garg, R., and Yusuf, S. (1995). Overview of randomized trials of angiotensin-converting enzyme inhibitors on mortality and morbidity in patients with heart failure. Collaborative Group on ACE Inhibitor Trials. *JAMA* 273, 1450–1456.
- [528] Breitmaier, E., and Jung, G. (2005). *Organische Chemie: Grundlagen, Stoffklassen, Reaktionen, Konzepte, Molekülstrukturen ; 129 Tabellen* (Georg Thieme Verlag).
- [529] Ahn, Y., Sanderson, B.W., Klein, O.D., and Krumlauf, R. (2010). Inhibition of Wnt signaling by Wise (Sostdc1) and negative feedback from Shh controls tooth number and patterning. *Dev. Camb. Engl.* 137, 3221–3231.
- [530] Murashima-Suginami, A., Takahashi, K., Sakata, T., Tsukamoto, H., Sugai, M., Yanagita, M., Shimizu, A., Sakurai, T., Slavkin, H.C., and Bessho, K. (2008). Enhanced BMP signaling results in supernumerary tooth formation in USAG-1 deficient mouse. *Biochem. Biophys. Res. Commun.* 369, 1012–1016.
- [531] Lintern, K.B., Guidato, S., Rowe, A., Saldanha, J.W., and Itasaki, N. (2009). Characterization of wise protein and its molecular mechanism to interact with both Wnt and BMP signals. *J. Biol. Chem.* 284, 23159–23168.
- [532] Henley, K.D., Gooding, K.A., Economides, A.N., and Gannon, M. (2012). Inactivation of the dual Bmp/Wnt inhibitor Sostdc1 enhances pancreatic islet function. *Am. J. Physiol. - Endocrinol. Metab.* 303, E752–E761.
- [533] Gadue, P., Huber, T.L., Paddison, P.J., and Keller, G.M. (2006). Wnt and TGF-beta signaling are required for the induction of an in vitro model of primitive streak formation using embryonic stem cells. *Proc. Natl. Acad. Sci. U. S. A.* 103, 16806–16811.
- [534] Naito, A.T., Shiojima, I., Akazawa, H., Hidaka, K., Morisaki, T., Kikuchi, A., and Komuro, I. (2006). Developmental stage-specific biphasic roles of Wnt/beta-catenin signaling in cardiomyogenesis and hematopoiesis. *Proc. Natl. Acad. Sci. U. S. A.* 103, 19812–19817.
- [535] Paige, S.L., Osugi, T., Afanasiev, O.K., Pabon, L., Reinecke, H., and Murry, C.E. (2010). Endogenous Wnt/ β -Catenin Signaling Is Required for Cardiac Differentiation in Human Embryonic Stem Cells. *PLoS ONE* 5, e11134.
- [536] Kattman, S.J., Witty, A.D., Gagliardi, M., Dubois, N.C., Niapour, M., Hotta, A., Ellis, J., and Keller, G. (2011). Stage-specific optimization of activin/nodal and BMP signaling promotes cardiac differentiation of mouse and human pluripotent stem cell lines. *Cell Stem Cell* 8, 228–240.

- [537] Nostro, M.C., Cheng, X., Keller, G.M., and Gadue, P. (2008). Wnt, activin, and BMP signaling regulate distinct stages in the developmental pathway from embryonic stem cells to blood. *Cell Stem Cell* 2, 60–71.
- [538] Schwartz, N.B. (1977). Regulation of chondroitin sulfate synthesis. Effect of beta-xylosides on synthesis of chondroitin sulfate proteoglycan, chondroitin sulfate chains, and core protein. *J. Biol. Chem.* 252, 6316–6321.
- [539] Kearns, A.E., Campbell, S.C., Westley, J., and Schwartz, N.B. (1991). Initiation of chondroitin sulfate biosynthesis: A kinetic analysis of UDP-D-xylose:core protein .beta.-D-xylosyltransferase. *Biochemistry (Mosc.)* 30, 7477–7483.
- [540] Pönighaus, C., Ambrosius, M., Casanova, J.C., Prante, C., Kuhn, J., Esko, J.D., Kleesiek, K., and Götting, C. (2007). Human xylosyltransferase II is involved in the biosynthesis of the uniform tetrasaccharide linkage region in chondroitin sulfate and heparan sulfate proteoglycans. *J. Biol. Chem.* 282, 5201–5206.
- [541] Götting, C., Sollberg, S., Kuhn, J., Weilke, C., Huerkamp, C., Brinkmann, T., Krieg, T., and Kleesiek, K. (1999). Serum xylosyltransferase: a new biochemical marker of the sclerotic process in systemic sclerosis. *J. Invest. Dermatol.* 112, 919–924.
- [542] Götting, C., Kuhn, J., Sollberg, S., Huerkamp, C., Brinkmann, T., Krieg, T., and Kleesiek, K. (2000). Elevated serum xylosyltransferase activity correlates with a high level of hyaluronate in patients with systemic sclerosis. *Acta Derm. Venereol.* 80, 60–61.
- [543] Schön, S., Schulz, V., Prante, C., Hendig, D., Szliska, C., Kuhn, J., Kleesiek, K., and Götting, C. (2006). Polymorphisms in the xylosyltransferase genes cause higher serum XT-I activity in patients with pseudoxanthoma elasticum (PXE) and are involved in a severe disease course. *J. Med. Genet.* 43, 745–749.
- [544] Ursell, P.C., Neill, C.A., Anderson, R.H., Ho, S.Y., Becker, A.E., and Gerlis, L.M. (1984). Endocardial fibroelastosis and hypoplasia of the left ventricle in neonates without significant aortic stenosis. *Br. Heart J.* 51, 492–497.
- [545] Lurie, P.R. (2010). Changing concepts of endocardial fibroelastosis. *Cardiol. Young* 20, 115–123.
- [546] Sellers, F.J., Keith, J.D., and Manning, J.A. (1964). The Diagnosis of Primary Endocardial Fibroelastosis. *Circulation* 29, 49–59.
- [547] Moller, J.H., Lucas, R.V., Adams, P., Anderson, R.C., Jorgens, J., and Edwards, J.E. (1964). Endocardial Fibroelastosis. A clinical and anatomic study of 47 patients with emphasis on its relationship to mitral insufficiency. *Circulation* 30, 759–782.
- [548] McElhinney, D.B., Vogel, M., Benson, C.B., Marshall, A.C., Wilkins-Haug, L.E., Silva, V., and Tworetzky, W. (2010). Assessment of left ventricular endocardial fibroelastosis in fetuses with aortic stenosis and evolving hypoplastic left heart syndrome. *Am. J. Cardiol.* 106, 1792–1797.
- [549] Friedman, K.G., Schidlow, D., Freud, L., Escobar-Diaz, M., and Tworetzky, W. (2014). Left ventricular diastolic function and characteristics in fetal aortic stenosis. *Am. J. Cardiol.* 114, 122–127.

- [550] Girirajan, S., Rosenfeld, J.A., Cooper, G.M., Antonacci, F., Siswara, P., Itsara, A., Vives, L., Walsh, T., McCarthy, S.E., Baker, C., et al. (2010). A recurrent 16p12.1 microdeletion supports a two-hit model for severe developmental delay. *Nat. Genet.* 42, 203–209.
- [551] Girirajan, S., Rosenfeld, J.A., Coe, B.P., Parikh, S., Friedman, N., Goldstein, A., Filipink, R.A., McConnell, J.S., Angle, B., Meschino, W.S., et al. (2012). Phenotypic heterogeneity of genomic disorders and rare copy-number variants. *N. Engl. J. Med.* 367, 1321–1331.
- [552] Lage, K., Greenway, S.C., Rosenfeld, J.A., Wakimoto, H., Gorham, J.M., Segrè, A.V., Roberts, A.E., Smoot, L.B., Pu, W.T., C. Pereira, A., et al. (2012). Genetic and environmental risk factors in congenital heart disease functionally converge in protein networks driving heart development. *Proc. Natl. Acad. Sci. U. S. A.* 109, 14035–14040.
- [553] Davies, P.F. (1995). Flow-mediated endothelial mechanotransduction. *Physiol. Rev.* 75, 519–560.
- [554] Topper, J.N., and Gimbrone Jr, M.A. (1999). Blood flow and vascular gene expression: fluid shear stress as a modulator of endothelial phenotype. *Mol. Med. Today* 5, 40–46.
- [555] Thornburg, K.L., and Louey, S. (2005). Fetal roots of cardiac disease. *Heart* 91, 867–868.
- [556] Ahmad, F., Seidman, J.G., and Seidman, C.E. (2005). The genetic basis for cardiac remodeling. *Annu. Rev. Genomics Hum. Genet.* 6, 185–216.
- [557] Knippers, R. (2006). *Molekulare Genetik* (Georg Thieme Verlag).

8 Anhang

Tabelle 30 Beschreibung verschiedener Operationsverfahren (mit Kürzel der Tabelle 1) und der Gewebeexzision im Rahmen dieser Operationen. Das Gewebe wurde für molekulargenetische und epigenetische Analysen in dieser Arbeit verwendet. Die Operationsverfahren wurden nach Angaben des Lehrbuchs von Apitz (2013) [13] beschrieben.

Operationsverfahren	Diagnose Patient (Alter bei Operation)	Beschreibung Operation und Gewebeexzision
Norwood-Operation (N)	HLHS (erste Lebensstage)	Beim HLHS kann die unterentwickelte linke Herzhälfte nicht die Systemperfusion aufrechterhalten. Durch die Norwood-Operation wird eine ungehinderte Systemperfusion ermöglicht, indem eine Anastomisierung (Verbindung) der durch einen <i>Patch</i> erweiterten hypoplastischen Aorta mit der Pulmonalarterie durchgeführt wird. Zur Sicherstellung der Lungenperfusion wird ein systemisch-pulmonaler Shunt angelegt. Mit der Norwood-Operation müssen Körper- und Lungenkreislauf parallelgeschaltet von der rechten Herzkammer mit Blut versorgt werden. Um einen adäquaten interatrialen Fluss sicher zu stellen, wird interatriales Septum exziiert (Atrioseptektomie), welches in dieser Arbeit für genetische, epigenetische und Transkriptom-Analysen verwendet wurde.
Hemifontan-Operation (H)	HLHS (4. bis 6. Lebensmonat)	Die Volumenbelastung der rechten Herzkammer wird durch eine Anastomisierung der oberen Hohlvene mit der Pulmonalarterie beseitigt (Hemifontan-Zirkulation). Der aortopulmonale Shunt ist nicht mehr notwendig und wird entfernt. Im Falle einer unzureichenden Atrioseptektomie bei der Norwood-Operation kann bei diesem OP-Schritt noch IAS-Gewebe entfernt werden, welches für die verschiedenen Analysen dieser Arbeit verwendet wurde.
totale kavopulmonale Anastomose (TCPC)	HLHS, TA (12. bis 18. Lebensmonat)	Die TCPC ermöglicht eine verbesserte Sauerstoffsättigung des Blutes. Die untere Hohlvene wird hierbei durch einen operativ geschaffenen Tunnel mit der Pulmonalarterie verbunden, sodass das gesamte venöse Blut zum Lungenkreislauf gelangt. Im Anschluss sind Körperkreislauf und Lungenkreislauf vollständig voneinander getrennt (Fontan-Zirkulation).
Aortopulmonale Shuntanlage (AP)	TA, (HLHS) (erste Lebensstage)	Aortopulmonale Shunt-Anlage, die die Lungenperfusion ermöglicht
Glenn-Anastomose (GL)	TA (3. bis 6. Lebensmonat)	Alternative chirurgische Technik zur Hemifontan-Operation (die obere Hohlvene wird direkt mit der rechten Pulmonalarterie verbunden).
AV-Kanal Korrektur (AV)	AVSD (4. bis 6. Lebensmonat)	Die Septumdefekte werden durch einen <i>Patch</i> (Flicken) verschlossen und die AV-Klappen rekonstruiert.

HLHS: Hypoplastisches Linksherzsyndrom; TA: Trikuspidalatresie; AVSD: atrioventrikulärer Septumdefekt; AV-Kanal: Atrioventrikularkanal; DA: *Ductus arteriosus*

Tabelle 31 Primer für die Sequenzierungen von PCR-Produkten

Primer Name [HGNC_fw/rv]	Primer-Sequenz [5'...3']	Position CHR (GRCh37/hg19) [CHR:Start-Ende]	Ampl.- Länge [bp]	Primer T _m [°C]
ABLIM2_fw	GCTGTGCTAAGGACACCCTACTCC	chr4:7986474- 7986756	283	60
ABLIM2_rv	CTGCGGATGGGGTGGTTTCAG			
ACE2_fw	GCCTGGCTAGGTGTATATTCTTATG	chrX:15582138- 15582475	338	65
ACE2_rv	GCCACTTACTTCTTCCGATCTC			
ACSL5_fw	GGATCACTGCAAATTTCTACACTAT	chr10:114171484- 114171972	489	60
ACSL5_rv	CATGAAGGCAGGAGCAAGTAC			
ANKRD11_fw1	CTGTCTCACCTCGGCCAGG	chr16:89346331- 89346625	295	65
ANKRD11_rv1	GCTCAGGCTCGAGCTCTGC			
ANKRD11_fw2	CTGGGCGGACGCCCTTCG	chr16:89346362- 89346593	232	65
ANKRD11_rv2	GGAGGCCTGGTCAGGAGG			
ANKRD11_fw3	TAGCTTCTGGACGGCAGC	chr16:89346325- 89346650	326	60
ANKRD11_rv3	CTGAGGGCTCAGGCTCGAG			
ASPG_fw	CTCCGCGCAGTCCCTGAGTC	chr14:104552029- 104552196	168	60
ASPG_rv	GACTCACCGCCGAGCTCACTC			
BHLHA15_fw1	GGTACCATTGCTGTCAGCAAGG	chr7:97841464- 97841744	281	65
BHLHA15_rv1	CAGACCCTTGCCGGCTC			
BHLHA15_fw2	GGTGACCACAGAGTGCCTGTG	chr7:97841533- 97841846	314	60
BHLHA15_rv2	CTGGATGCTGCTGTCACGAC			
BHLHA15_fw3	GTGAACAGAAGGACCTCCAGTCG	chr7:97841419- 97841744	326	65
BHLHA15_rv3	CAGACCCTTGCCGGCC			
C15orf62_fw	CCAGCACAGCAGGTGGG	chr15:41062800- 41063231	432	65
C15orf62_rv	CTCACCCCATTTAGAAGCTGG			
CCDC125_fw	GATGGAAAGAAGTTTAGCCCTCC	chr5:68616025- 68616211	187	65
CCDC125_rv	CACAGGAGAAAGAAAGGAGAGG			
CREBBP_fw	CAGGCTTACAAGGGCAG	chr16:3900316- 3900498	183	65
CREBBP_rv	CTGTGCGGTGTTCACTCC			
CTNNB1_fw	GGGAACCTCGGGTATATAATG	chr3:41277113- 41277334	222	55
CTNNB1_rv	CTGCACAAACAATGGAATGG			
DNAH3_fw	GACGTTCCAGGAAGCAGAC	chr16:21151756- 21151968	213	65
DNAH3_rv	CATTTCCAGGTGATTCTGC			
DNAH8_fw	GTTATGGAGGAATATGATAGTGA	chr6:38854510- 38854699	190	60
DNAH8_rv	CATTTGACCACCTCCAGG			
DOCK1_var1_fw	CTGTTTCCAGTGACCTTCTGG	chr10:128835973- 128836168	196	55
DOCK1_var1_rv	CCTTTGAAGTCACTGCCTAC			

Primer Name [HGNC_fw/rv]	Primer-Sequenz [5'...3']	Position CHR (GRCh37/hg19) [CHR:Start-Ende]	Ampl.-Länge [bp]	Primer T _m [°C]
DOCK1_var2_fw	GTGTCTGTGGGCTGGAAAG	chr10:129237231-129237417	187	60
DOCK1_var2_rv	CTCTTGATGTTTGCTGTTCC			
FAM65A_fw	GGTGGTTCTGAAATGCCCTCTC	chr16:67575737-67575943	207	60
FAM65A_rv	CCTGCGGAAGGCAACTTGG			
FBXW7_fw	GGTGGCATCAATACTTTTTTGC	chr4:153332737-153333143	407	65
FBXW7_rv	CTGGGAATCATTTTGGCCT			
GALE_fw	CCTCAGCCTGCCTAGCACC	chr1:24122636-24122882	247	60
GALE_rv	CTCACACATCCTGTCCAGCC			
GDPD5_fw	CATTCTGGTGGGAGGCAGC	chr11:75173769-75174005	237	60
GDPD5_rv	GTCCCAGCTACTGGATACCC			
GPD2_fw	GATGATTTCTCATCAGGGACC	chr2:157367334-157367594	261	65
GPD2_rv	CTGACACACCTCTCCTGAAAG			
GRB7_fw	CATCCTGGTGGTAGAGGGG	chr17:37900255-37900607	353	65
GRB7_rv	CAAGGATGTAGGGAGAGATGG			
HEPHL1_fw	AACAGTTTACATATTTGAAGCTCCC	chr11:93819068-93819466	399	65
HEPHL1_rv	CTACACCACTTGAAGTGCACAAAG			
HIST1H2AG_fw	CCAAGTGGCCATCCGCAAC	chr6:27101102-27101312	211	60
HIST1H2AG_rv	GGTGGCCCTGAAAAGAGCCG			
HSPA9_fw	GGAGTGTTCCTGCAAGGAC	chr5:137903111-137903559	449	60
HSPA9_rv	CATCCAGGTGAGAAATTTACTGC			
IFT88_fw	GAAAGGATTTTGCACAAGCC	chr13:21212472-21212645	174	55
IFT88_rv	GTGCTTCAGTACAAGAAGAATC			
IMPAD1_fw	GATGCTGGCTGTGTCAGTGC	chr8:57905770-57905950	181	60
IMPAD1_rv	GAAGGCGGTCTTGAGCAG			
KBTBD3_fw	CATTCACAAAAGTGGGAAAATGG	chr11:105924088-105924522	435	65
KBTBD3_rv	CCTCTGGGTAATGGGCTAACA			
KCND1_fw	GAGTGCATCCAGGCCTTC	chrX:48826133-48826354	222	65
KCND1_rv	GCTCGTGTGTGGATTCTCG			
KLF17_fw	CCTTCCACTGAGGCCAG	chr1:44595532-44595976	445	65
KLF17_rv	AGCTTGAAGATGAAAGGCTC			
KLHL8_fw	CAGGTGGGAGCTTGTGGATC	chr4:88084595-88084797	203	60
KLHL8_rv	AGGTGGTCATATCAAAATCTTGG			
LCLAT1_fw	GCTGTCCATGATATCACTGTGGC	chr2:30862997-30863249	253	60
LCLAT1_rv	TGCAAGGTGGAATGACACTCTG			

Primer Name [HGNC_fw/rv]	Primer-Sequenz [5'...3']	Position CHR (GRCh37/hg19) [CHR:Start-Ende]	Ampl.-Länge [bp]	Primer T _m [°C]
LMNB2_fw	GGTGTGGAAGGGCCAGAG	chr19:2431402-2431853	452	60
LMNB2_rv	GCCAAGGCTGGCTTCAC			
LPXN_fw	GAACGCTCTGCCATGGGTGTG	chr11:58294903-58295111	209	60
LPXN_rv	GAAGAGGCTATGGATCAGTTGGC			
LRIT2_fw	GCTGAGACTGGAGGGGAAC	chr10:85984460-85984660	201	65
LRIT2_rv	CAGTTCAGGAAGACACTCTTGG			
LYPD3_fw	CGAGGTTACTCATCCTGGGCTC	chr19:43969636-43969811	176	60
LYPD3_rv	CAGACCCACCTCCGCGAAG			
MADD_fw	GGTAGAGGAGAAGTCAAGTGC	chr11:47297486-47297669	184	55
MADD_rv	CACTAGGGTGAATCGAGATGG			
MAPK4_fw	GCTGCCTCTTGCTTTAATTTGAG	chr18:48184593-48184996	404	65
MAPK4_rv	CCACTTGCATCTTAACAGAAGC			
MIIP_fw	TTGGGGGACAGAGCCTG	chr1:12089978-12090317	340	60
MIIP_rv	GGCCCAGGCTGTGCTC			
MYH7B_fw	GCACTTGCTCTGACCTGGTG	chr20:33584065-33584485	421	60
MYH7B_rv	CCTCCAGCTTGCCTTG			
NOTCH1_fw	GATAGCACCTAAGGTGGAGCC	chr9:139417542-139417754	213	55
NOTCH1_rv	GTGGGCAGTGGCAGATGTAG			
NR2F2_fw	GAGGATGCTCTTCAGCGCCGTC	chr15:96877516-96877996	481	65
NR2F2_rv	CCTGGTAGTGCCAGCAGC			
NR2F2_fw1	GACCTGCAGATCACGGACC	chr15:96877565-96878030	466	65
NR2F2_rv1	CAATCCAGCCTTCTCTCTC			
NUP37_fw	TCACTTACAAAGGCTAAAAATTGG	chr12:102494700-102494931	232	65
NUP37_rv	TAATGATGACCATAAAAAGCC			
OR1S2_fw	CTTTCATGCGGGCCAGTTTC	chr11:103070591-103070904	251	60
OR1S2_rv	GGAGAAGAAGATGAGTACAAAGGG			
PABPC3_fw	GCATGCTCTGGACACCATG	chr13:25670534-25670979	446	65
PABPC3_rv	CGAACTTGCCAAAGAGATCC			
PARK2_fw	CTCCAGGATTACAGAAATTGGTC	chr6:162206610-162207028	419	60
PARK2_rv	CAGTTC AACACAATTCCTTCATTC			
PDZD2_fw	GTTTAAATTGGGTCCACACC	chr5:32010263-32010714	452	60
PDZD2_rv	GCACTGGTGTCTTTTCCCATC			
PNMA1_fw	CTGTTGGTAGAATTTGGCGTC	chr14:74180178-74180378	201	65
PNMA1_rv	GAACATTCTCCAAGCATTTCG			
PRF1_fw	CTGCACCCTGACTTCAAGAG	chr10:72358452-72358912	461	65
PRF1_rv	GTGTAGTCCACCAGGCCAG			

Primer Name [HGNC_fw/rv]	Primer-Sequenz [5'...3']	Position CHR (GRCh37/hg19) [CHR:Start-Ende]	Ampl.-Länge [bp]	Primer T _m [°C]
PTCH2_fw	GCATGCTCCCTGTACAGTTCC	chr1:45288783-45289210	428	60
PTCH2_rv	CACGAGTCCATGGAGGAGG			
RADIL_fw	CAGCCCAGGCATCACCCCTC	chr7:4854909-4855117	209	65
RADIL_rv	CTGGATGAGCTGGGCCCTG			
RNF207_fw	CTGAGTGCTTTGGCCCAAGTGG	chr1:6279209-6279474	266	60
RNF207_rv	CCATTTGCTAAGTCGGGTGTTCC			
ROR2_fw	CCAGGCTGTGGCCATCA	chr9:94487101-94487273	173	60
ROR2_rv	CGTGCGAACAGTAGCTGAAG			
SCYL3_fw	GATCTCGCCAATCTTGGAG	chr1:169824853-169825061	209	60
SCYL3_rv	CTCCCAAACACCCATAAGC			
SPATA2_fw	CAAGCCCTCGAGGTCAGTG	chr20:48522505-48522975	471	65
SPATA2_rv	GGACAGGTGAGCAGGCTG			
SPERT_fw	TGCAGCTCAGCGACGAG	chr13:46287522-46287850	329	60
SPERT_rv	CTTACCACCTCCAGGGAC			
SMO_fw	CCCATCCCTGACTGTGAGATC	chr7:128850244-128850445	202	60
SMO_rv	GTCCCAAGGCTGAGGGTG			
SNX4_fw	CTCTCTCCAATTCCTGCTGAAG	chr3:125216068-125216341	274	60
SNX4_rv	ATTGTAAGACGATTTCCAACATTC			
SNX13_fw	CTGGAGGGATCAGAGGAAATC	chr7:17873562-17873976	415	65
SNX13_rv	ACAGTTTATATTTGGACTCCAACAAC			
SOSTDC1_fw	GCCTGCAAGTGCAAGAGGTAC	chr7:16502090-16502310	221	60
SOSTDC1_rv	GCAGTGGCAGGCTTGAGTC			
STEAP2_fw	CGGCTATCATGTGGTCATAGG	chr7:89854552-89854784	233	60
STEAP2_rv	CATTGGATTCTGGGTAAGTGG			
TGDS_fw	GAGTGCTGTATGAAACCTAGGCC	chr13:95242983-95243245	263	60
TGDS_rv	CACAGTCACCGTGGGCTTG			
TNKS1BP1_fw	GTGACCTGTGACCCAGACTCT	chr11:57075782-57076221	440	60
TNKS1BP1_rv	GTTTCTTCCCTCCTCAGAAAGG			
TNNC2_fw	GTTTGATGCTGATGGTGGTGGG	chr20:44453053-44453255	203	60
TNNC2_rv	CAGCAGGTGGCAGACTGAGC			
USP24_fw	GCTGGAATAATGGGCCCTTTG	chr1:55560931-55561156	226	60
USP24_rv	GAGAGATCCTCAGCTTCCCTGCC			
ZNF418_fw	CATCAACGAGGTCACACTAGTG	chr19:58438255-58438471	217	60
ZNF418_rv	CACATTCTCCACACTCGTAAGG			

Primer Name [HGNC_fw/rv]: HGNC: Genname nach HUGO Gene Nomenclature Committee (HGNC), fw: Vorwärts-Primer, rv: Rückwärts-Primer. Primer-Sequenz [5'...3']: Primersequenz in 5' nach 3' Richtung. Position CHR (GRCh37/hg19) [CHR:Start-Ende]: chromosomale Position des Amplifikats nach GRCh37/hg19. Ampl.-Länge [bp]: Amplifikatlänge in Basenpaaren (bp). Primer T_m [°C]: Primer-Schmelztemperatur (T_m) in °C

Tabelle 32 Primer für die Bisulfit-Pyrosequenzierung

Primer Name [HGNC_TargetID__F/R/S]	Primer-Sequenz [5'...3']	Position CHR (GRCh37/hg19) [CHR:Start-Ende]	Ampl.- Länge [bp]	Primer T _m [°C]	Ziel-CpG [TargetID]
CTNNB1_cg09678212_F	AGTGATAAGTGGGAATTAGAT AAAAATGG	chr3:41240037- 41240282	246	60	cg09678212
CTNNB1_cg09678212_R	TAAATCCCACCTCTCAACAA ACTATTCA				
CTNNB1_cg09678212_S	GTTTTGGAGTTAATTTATTTT TATT				
NR2F1_cg04192879_F	GTGAGGAGTTAGTATTTTAA TTAGTTTAGT	chr5:92929342- 92929612	271	60	cg04192879
NR2F1_cg04192879_R	ATCCTTAACCCAAATAAATC CTCT				
NR2F1_cg04192879_S	TTTTTAGTTATAGAGTAGTT TTTTTT				
NR2F1_cg25449484_F	GGTTTATATAGAGAGTTTGT AGGAGAAGT	chr5:92929290- 92929518	229	60	cg25449484
NR2F1_cg25449484_R	ATTAATACTACTCCCAAACA ATAACATA				
NR2F1_cg25449484_S	AGTAGTAGTGAGTATTGGA GGAGTA				
NR2F1-AS1_cg11080549_F	AAGGTTTTGGTAGTATGAAT TGAGTAG	chr5:92909859- 92910087	229	60	cg11080549
NR2F1-AS1_cg11080549_R	AAAAAACTTACCTTTTACCC CTCTA				
NR2F1-AS1_cg11080549_S	AATGGTTTTTTTTGTAGGTT TT				
NR2F1-AS1_cg24579887_F	GGATATTGGGAGATTGTTAT TTTAAGAG	chr5:92905584- 92905823	240	60	cg24579887
NR2F1-AS1_cg24579887_R	TCTCCACATTACCTACTCAC ACTTA				
NR2F1-AS1_cg24579887_S	ACAAACAATCCTCCC				
PDXDC1_cg03245889_F	TGTGTGTGGTTAGGATTTTG AATAG	chr16:15127041- 15127259	219	60	cg03245889
PDXDC1_cg03245889_R	AACCCTCCAATTCTCCTCTA TCTCCC				
PDXDC1_cg03245889_S	AGTTTGTTTAGTTATTTTTGT AGG				
PDXDC1_cg06978461_F	TTTGTAGGTTTTGAGTATAA GAGTATGA	chr16:15127128- 15127333	206	60	cg06978461
PDXDC1_cg06978461_R	ACTACTTTCCTCCTAACCCC TACTATTTAA				
PDXDC1_cg06978461_S	GTTTTGAGTATAAGAGTATG AAG				
STEAP2_cg02206373_F	TTTTGTTATTTTGGGTTGAT GTAGGG	chr7:89841709- 89841869	161	60	cg02206373
STEAP2_cg02206373_R	AACAATCAAAAAAAAAACCA ATTACTTTC				
STEAP2_cg02206373_S	GGGTTGATGTAGGGTA				
XYLT1_cg05057352_F	GTTAGGAGTTGGGTTTGT GTTA	chr16:17499183- 17499418	236	60	cg05057352
XYLT1_cg05057352_R	AACTACCAAAACCCAATAAA AAAA				
XYLT1_cg05057352_S	TTTAATTTTTTAGGGTTTTTA GTG				

Primer Name [HGNC_TargetID__F/R/S]	Primer-Sequenz [5'...3']	Position CHR (GRCh37/hg19) [CHR:Start-Ende]	Ampl.- Länge [bp]	Primer T _m [°C]	Ziel-CpG [TargetID]
XYLT1_cg09770579_F	TGGGTAGAATTATTTTAAAT GTTATTATGG	chr16:17407321- 17407601	281	60	cg09770579
XYLT1_cg09770579_R	ACTCCTTAAAAAAAATATT CCCTATACT				
XYLT1_cg09770579_S	AGTTATATAGATAGTATTAT AGTAG				

Primer Name [HGNC_TargetID_F/R/S]: HGNC: Gennamen nach HUGO Gene Nomenclature Committee (HGNC); [TargetID]: Ziel-CpG, zu analysierendes CG-Dinukleotid mit Angabe der Illumina TargetID; F: Vorwärts-Primer, R: Rückwärts-Primer, S: Sequenzierprimer. Primer-Sequenz [5'...3']: Primersequenz in 5' nach 3' Richtung. Position CHR (GRCh37/hg19), [CHR:Start-Ende]: chromosomale Position des Amplifikats nach GRCh37/hg19. Ampl.-Länge [bp]: Amplifikatlänge in Basenpaaren (bp). Primer T_m [°C]: Primer-Schmelztemperatur (T_m) in °C. Ziel-CpG [TargetID]: zu analysierendes CG-Dinukleotid mit Angabe der Illumina TargetID.

Tabelle 33 R-Befehle zum R-Package RnBeads anhand eines Beispiels. A) Befehle zur RnBeads (Qualitäts-)Analyse, B) Befehle zur Normalisierung

R-Befehle (je > - Zeichen)	Kommentar
<pre>>library(RnBeads) >Sys.setenv(R_GSCMD="C:/Program Files/gs/g9.07/bin/gswin64c.exe") >rnb.options(disk.dump.big.matrices = TRUE) >data.dir <- "C:/Dateipfad/HLHS" >idat.dir <- file.path(data.dir, "idats_RnBeads_HLHS") >sample.annotation <- file.path(data.dir, "Annotation_HLHS.csv") >analysis.dir <- "C:/Dateipfad/HLHS/Normalisierung" >report.dir <- file.path(analysis.dir, "reports") >data.source <- c(idat.dir, sample.annotation) >rnb.options(identifiers.column="Sample_ID")</pre>	<p>Laden RnBeads <i>Package</i></p> <p>Dateipfad Import IDAT-Ordner Annotations-Tabelle Dateipfad Export Ergebnisse</p> <p>Datenquelle Definition Probenspalte</p>
<p>A) RnBeads Analyse</p> <pre>>analysis.dir <- "C:/Users/hoff-k/Dokumente_KHoff/UK-SH/450K-Array/MethylierungsArray_HLHS/RnBeads/analysis" >report.dir <- file.path(analysis.dir, "reports") >rnb.options(filtering.sex.chromosomes.removal = TRUE, identifiers.column="Sample_ID") >rnb.run.analysis(dir.reports = report.dir, sample.sheet = sample.annotation, data.dir= idat.dir, data.type = "idat.dir")</pre>	<p>A) RnBeads-Analyse Dateipfad Export Ergebnisse</p> <p>Optionen Filterung und Definition Probenspalte RnBeads-Analyse (Qualitätskontrolle)</p>
<p>B) Normalisierung</p> <pre>>result<-rnb.run.loading(data.source = data.source, data.type = "idat.dir", dir.reports=report.dir) >result.norm<-rnb.run.normalization(result[[1]], dir.reports = report.dir) >rnb.execute.export.csv(result.norm\$rnb.set, "C:/Dateipfad/HLHS/Normalisierung")</pre>	<p>B) Normalisierung Hochladen der Import-Dateien Normalisierung Export Exceltabelle (Normal-Tabelle)</p>

Tabelle 34 R-Befehle unter Verwendung der R-basierten Software Bioconductor (BiocLite-Tool) zur Erstellung einer Import-Datei (hier „Import_EpiDesigner.csv“) anhand einer Liste der zu analysierenden CpGs (hier „CpG-Liste.txt“) für die EpiDesigner-Software

R-Befehle (je > - Zeichen)	Kommentar
<pre>>source("http://bioconductor.org/biocLite.R") >biocLite("Biostrings") >biocLite("BSgenome") >biocLite("BSgenome.Hsapiens.UCSC.hg19") >library(Biostrings) >library(BSgenome) >library("BSgenome.Hsapiens.UCSC.hg19") >manifest<- read.csv("C:/Dateipfad1/HumanMethylation450_15017482_v.1.2.csv", header=T,skip=7,sep=",") >AnalyseCpG<-read.delim("C:/Dateipfad2/CpG-Liste.txt ") >AnalyseCpG_m<-manifest[manifest\$Name %in% AnalyseCpG[,1],] >pos1<- AnalyseCpG_m\$MAPINFO-200 >pos2<- AnalyseCpG_m\$MAPINFO+200 >chr<-paste("chr", AnalyseCpG_m\$CHR,sep="") >GenomSequenz<-getSeq(Hsapiens, names= chr, start= pos1, end= pos2,as.character = TRUE) >CpGID<- AnalyseCpG_m\$Name >UCSCannotation<-paste(">",chr,":",pos1,"-",pos2,sep="") >TabelleEpiDesign <- data.frame(CpGID,UCSCannotation,GenomSequenz) >write.table(TabelleEpiDesign,"Import_EpiDesigner.csv",sep=",",row.names =F,quote=F)</pre>	<p>Installieren &Laden von BiocLite Package und Funktionen</p> <p>GRCh37/hg19</p> <p>Manifest-Datei 450K-Array⁴⁹ einladen</p> <p>CpG-Liste einladen Verbinden mit CpGs Manifest 200 bp <i>down-/upstream</i> von CpG extrahieren CHR extrahieren/definieren Genomische Sequenz nach pos1/pos2 (s.o.) extrahieren CpG-Bezeichnung extrahieren Annotation extrahieren/definieren Tabelle erstellen mit CpG- Bezeichnung, Annotation, Genom. Sequenz (402 bp-Sequenz) Import_EpiDesigner-Tabelle exportieren</p>

Tabelle 35 R-Befehle unter Verwendung des R-Packages MassArray zur *in silico* Assay-Vorhersage für die Epi typer-Technik der zuvor ermittelten Amplikon-Sequenzen

R-Befehle (je > - Zeichen)	Kommentar
<pre>>source("http://bioconductor.org/biocLite.R") >biocLite("MassArray") >library(MassArray) >ampliconPrediction("GGGACAGGGGTCTGATTCAC>CTACCAACCCTCTCCC TCCACTGGCCACGCATTCTCTCCCGCAGAG...CGCGCATCTGGTCCCTGGCC CACACAGCCACCGCCGCCGCCGCC<GCCACCTCCCTGAGCCAGACTG AGTTTC",lower.threshold = 1500, upper.threshold = 7000) ... = weitere 300 bp Sequenz wurden hier zur Vereinfachung nicht dargestellt</pre>	<p>Installieren &Laden von MassArray Package</p> <p>("Vorwärts-Primer>...zu analysierende Sequenz... <Rückwärts-Primer" Definition messbarer Bereich (Mass Window)</p>

Tabelle 36 R-Befehle zur Erstellung einer Korrelationsmatrix anhand eines Beispiels

R-Befehle (je > - Zeichen)	Kommentar
<pre>>data=read.csv("HLHS.csv", header=TRUE) >w=array('e',dim=c(36,36)) >for(t in 1:36){ for(i in 1:36){ r=cor(data[,t],data[,i],method="pearson", use="everything") w[t,i]=r}} >write.table(w, file="HLHS_CpGKorrelationsmatrix.txt", sep="/t", row.names=TRUE, col.names=TRUE, quote=FALSE, dec=".", append=FALSE)</pre>	<p>Importieren der Tabelle Definition Array-Matrix Schleife Korrelationsanalyse von jeder Zeile (CpG-Locus) von 36 Spalten (Proben)</p> <p>Exportieren der Korrelationsmatrixtabelle (wird anschließend bedingt formatiert)</p>

⁴⁹ Manifest-Datei entspricht der Annotationstabelle für den 450K-Array mit zusätzlichen Informationen zu den CpG Loci, wie UCSC_RefGene_Name, Mapinfo, etc.. Diese Datei entspricht hier der Datei, die in der GS-Software zur 450K-Methylierungsanalyse verwendet wurde (siehe 2.2.2.2)

Tabelle 37 Übersicht über die 21 Kandidatengene, die mittels Sanger-Sequenzierung zur Untersuchung des Vererbungsmodells von Familie 5 analysiert wurden.

Gen (HGNC)	NM-Nummer	Position CHR (GRCh37/hg19)	VEP	Mutation (Konsequenz Proteinebene)	Index-1	Index-2	Mutter	Vater	Bruder	VM
<i>MIIP</i>	021933.3	chr1:12090104	Miss	c.1258A>T(p.Ile289Phe)	c.1258A>T	WT	c.1258A>T	WT	c.1258A>T	nein
<i>KLF17</i>	173484.3	chr1:44595787	Miss	c.902C>T(p.Pro282Ser)	c.902C>T	c.902C>T	c.902C>T	WT	c.902C>T	nein
<i>PTCH2</i>	001166292.1	chr1:45289033	Miss	c.3151G>A(p.Arg1047Trp)	c.3151G>A	WT	c.3151G>A	WT	c.3151G>A	nein
<i>ACSL5</i>	203380.1	chr10:114171726	Miss	c.1244A>T(p.Arg313Trp)	c.1244A>T	c.1244A>T	c.1244A>T	WT	c.1244A>T	nein
<i>PRF1</i>	005041.4	chr10:72358720	Stop	c.757G>T (p.Glu253*)	c.757G>T	c.757G>T	c.757G>T	WT	WT	ja
<i>KBTBD3</i>	152433.3	chr11:105924310	Miss	c.1545G>C(p.Ala369Gly)	c.1545G>C	WT	WT	c.1545G>C	WT	nein
<i>TNKS1BP1</i>	033396.2	chr11:57075987	Miss	c.4198G>A(p.Gly1400Arg)	c.4198G>A	c.4198G>A	c.4198G>A	WT	WT	ja
<i>HEPBL1</i>	001098672.1	chr11:93819308	Miss	c.2190T>C(p.Leu678Pro)	c.2190T>C	WT	WT	c.2190T>C	WT	nein
<i>PABPC3</i>	030979.2	chr13:25670755	Miss	c.480A>G(p.Tyr140Cys)	c.480A>G	WT	WT	c.480A>G	WT	nein
<i>SPERT</i>	152719.1	chr13:46287725	Miss	c.645C>G(p.Leu189Val)	WT	c.480A>G	c.480A>G	WT	c.645C>G	nein
<i>C15orf62</i>	001130448.2	chr15:41063028	Miss	c.335C>G (p.Ser112Trp)	c.335C>G	c.335C>G	c.335C>G	WT	WT	ja
<i>GRB7</i>	001242442.1	chr17:37900401	Miss	c.946C>T(p.Arg271Cys)	c.946C>T	WT	c.946C>T	WT	c.946C>T	nein
<i>MAPK4</i>	002747.3	chr18:48255706	Miss	c.2246A>T(p.Met416Leu)	c.2246A>T	WT	c.2246A>T	WT	WT	nein
<i>LMNB2</i>	032737.3	chr19:2431617	Miss	c.1785G>A(p.Arg584Cys)	WT	c.1785G>A	c.1785G>A	WT	c.1785G>A	nein
<i>MYH7B</i>	020884.3	chr20:33584231	Miss	c.3469C>T(p.Ala1051Val)	c.3469C>T	WT	c.3469C>T	WT	c.3469C>T	nein
<i>SPATA2</i>	001135773.1	chr20:48522730	Miss	c.1339G>A(p.Ser330Phe)	WT	c.1339G>A	WT	c.1339G>A	c.1339G>A	nein
<i>FBXW7</i>	033632.3	chr4:153332916	Stop	c.202G>A(p.Arg14*)	WT	c.202G>A	WT	c.202G>A	WT	nein
<i>HSPA9</i>	178140.2	chr5:137903382	Miss	c.1613C>A(p.Ser430Pro)	c.1613C>A	WT	WT	c.1613C>A	WT	nein
<i>PDZD2</i>	004134.6	chr5:32010469	Miss	c.876T>C(p.Ala189Ser)	c.876T>C	c.876T>C	WT	c.876T>C	c.876T>C	nein
<i>PARK2</i>	013988.2	chr6:162206827	Miss	c.535A>G(p.Leu134Pro)	c.535A>G	WT	WT	c.535A>G	c.535A>G	nein
<i>SNX13</i>	015132.4	chr7:17873740	Miss	c.1745C>T(p.Arg511His)	WT	c.1745C>T	WT	c.1745C>T	c.1745C>T	nein

Position CHR: chromosomale Position der Mutation; VEP: *Variant Effect Predictor* (VEP) Konsequenz (1000G-rel77); Miss: Missense Variante, Stop: Stop Variante. Analytierte DNA-Proben der Familie Familie 5 (siehe 2.1.1.1, Tabelle 2, Stammbaum Abbildung 18 A) aus der Familien-Analyse in Abschnitt 3.1.2.2 sind in den Spalten rechts dargestellt, mit den beiden Patienten mit HLHS, Indexpatient-1 (HLHS_130_Index-1) und Indexpatient-2 (HLHS_130_Index-2), der phänotypisch gesunden Mutter (HLHS_130_Mutter), dem gesunden Vater (HLHS_130_Vater) und dem gesunden Geschwisterkind (HLHS_130_Bruder). Entsprechend des wahrscheinlichsten Vererbungsmodells „autosomal dominant mit reduzierter Penetranz“ wurde eine Vererbung der Mutationen über die Generationen mütterlicherseits vermutet. Bei Auftreten der Mutation in Mutter und beiden Indexpatienten, jedoch nicht bei Vater und gesundem Geschwisterkind, lagen die detektierten Mutationen nach dem genannten Vererbungsmodell vor. Dies ist mit der Spalte „VM“ (Vererbungsmodell) gekennzeichnet: ja = Modell bestätigt, nein = Modell nicht bestätigt. Die phänotypisch gesunde Mutter, sowie die Geschwister mit HLHS, Indexpatient-2 und Indexpatient-1, zeigten jeweils eine heterozygote Ausprägung der Mutationen in *C15orf62* (c.335C>G), *TNKS1BP1* (c.4198G>A) und *PRF1* (c.757G>T). WT: Mutation nicht vorhanden (WT-Sequenz).

Tabelle 38 Ergebnistabelle der kombinierten Keimbahn- und somatischen Analyse aus Abschnitt 3.2.1.2. Übersicht über die 11 heterozygoten Genvarianten, die nach Varbank-Filteroberfläche als „IAS“-spezifisch annotiert wurden und anschließend visuell hinsichtlich der *Reads Coverage* untersucht wurden. Für die IAS-Gewebeprobe und pB-Probe des Indexpatienten HLHS_4, sowie für pB-Proben von dessen Eltern (Familie_4, siehe 2.2.1.7 (Tabelle 13)) wurden die *Reads* der WES-Analyse überprüft und die Mutationen der Gene *ANKRD11* und *BHLHA15* mittels Sanger-Sequenzierung analysiert.

Gen HGNC	Mutation (Konsequenz)	Position CHR (GRCh37/hg19)	PPH	ExAC [AF]	Index IAS					Index pB					Mutter pB					Vater pB					Normalkontrolle	Auswertung
					MAF [%]	CV	VA	Mut. WES	Mut. Sanger	MAF [%]	CV	VA	Mut. WES	Mut. Sanger	MAF [%]	CV	VA	Mut. WES	Mut. Sanger	MAF [%]	CV	VA	Mut. WES	Mut. Sanger		
<i>ANKRD11</i>	c.6464C>A (p.P2155H)	16:89346486	p.d.	-	43	7	ok	ja	nein	0	12	ok	nein	nein	0	10	ok	nein	nein	0	18	ok	nein	nein	nein	potenziell somatisch
<i>BHLHA15</i>	c.6G>T (p.K2N)	7:97841627	p.d.	-	43	7	ok	ja	nein	0	3	↓	nein	nein	0	1	↓	nein	nein	0	6	ok	nein	nein	nein	potenziell somatisch
<i>COL9A3</i>	c.90_99delCCCCGGCCC (p.P34_P36del)	20:61448930-61448939	-	-	0.5	10	ok	ja	n.t.	0	17	ok	nein	n.t.	-	0	n.c.	-	n.t.	0	1	↓	nein	n.t.	n.t.	geringe/keine CV
<i>FAM98B</i>	c.1249G>T (p.G417C)	15:38776807	-	-	0.3	13	ok	ja	n.t.	0	14	ok	nein	n.t.	0	11	ok	nein	n.t.	-	0	n.c.	-	n.t.	n.t.	geringe/keine CV
<i>MPLKIP</i>	c.278C>A (p.S93Y)	7:40173889	b	0.00038	0.5	29	ok	ja	n.t.	-	1	↓	ja	n.t.	0	1	↓	nein	n.t.	0	1	↓	nein	n.t.	n.t.	geringe/keine CV auch in Keimbahn
<i>SAC3D1</i>	c.266G>A (p.S89N)	11:64809030	-	-	0.6	7	ok	ja	n.t.	-	1	↓	ja	n.t.	0	24	ok	nein	n.t.	-	0	n.c.	-	n.t.	n.t.	geringe/keine CV auch in Keimbahn
<i>MRPL55</i>	c.62C>T (p.A21V)	1:228295750	-	-	0.3	45	ok	ja	n.t.	-	0	n.c.	-	n.t.	-	0	n.c.	-	n.t.	-	1	↓	ja	n.t.	n.t.	auch in Keimbahn
<i>POC1B-GALNT4</i>	c.184C>T (p.P62S)	12:89919749	-	0.00054	0.4	8	ok	ja	n.t.	42	22	ok	ja	n.t.	-	0	n.c.	-	n.t.	-	1	↓	ja	n.t.	n.t.	auch in Keimbahn
<i>WDR90</i>	c.1837A>G (p.T613A)	16:705691	p.d.	-	0.6	51	ok	ja	n.t.	38	17	ok	ja	n.t.	42	23	ok	ja	n.t.	0	21	ok	nein	n.t.	n.t.	auch in Keimbahn
<i>KCNS2</i>	c.25G>A (p.V9M)	8:99440232	b	4x10 ⁻⁵	0.4	41	ok	ja	n.t.	44	24	ok	ja	n.t.	0	21	ok	nein	n.t.	0	27	ok	ja	n.t.	n.t.	auch in Keimbahn
<i>C9orf114</i>	c.35C>G (p.P12R)	9:131592025	p.d.	5,5x10 ⁻⁵	0.4	45	ok	ja	n.t.	41	21	ok	ja	n.t.	-	1	↓	ja	n.t.	0	22	ok	nein	n.t.	n.t.	auch in Keimbahn

Proben Index IAS, Index pB, Mutter pB und Vater pB entsprechen den DNA-Proben der Familie_4 (siehe 2.2.1.7 (Tabelle 13)). Zusätzlich wurden die Mutationen in einer unabhängigen DNA-Normalkontrolle analysiert. Mutation (Konsequenz): Nomenklatur Mutation in kodierender Sequenz und Konsequenz auf Proteinebene in Klammern; Position CHR: chromosomale Position (nach GRCh37/hg19); PPH: Polyphen-2; ExAC AF: Allelfrequenz nach ExAC (Version 0.3, Feb. 2015); p.d. = *probably damaging*; b = *benign*. Pro Probe sind die Anzahl Reads (CV) sowie die visuelle Reads-Analyse (VA) aufgeführt. MAF: mutierte Allelfrequenz [%]; CV: *Reads Coverage* (Anzahl *Reads*); VA: Visuelle *Reads*-Analyse; Mut. WES: Mutation in WES-Analyse vorhanden?; Mut. Sanger: Mutation in Sanger-Sequenzierung vorhanden?; ja = Mutation detektiert; nein = Mutation nicht detektiert; ok = mehr als 6 *Reads*; sowohl Vorwärts- und Rückwärts-Sequenzen vorhanden; ↓ = weniger als 6 *Reads*. n.c. = nicht gecovered; n.t. = nicht getestet aufgrund Auswertungskriterien; Auswertung: Bewertung der Mutationen anhand der VA hinsichtlich potenziell somatischer Mutationen (nur im IAS-Gewebe, nicht in pB-Proben). Geringe/keine CV: mind. zwei Proben zeigen eine geringe oder keine *Coverage*; auch in Keimbahn: die Mutation ließ sich in der VA der WES-Analyse auch in der Keimbahn (pB-Probe) detektieren.

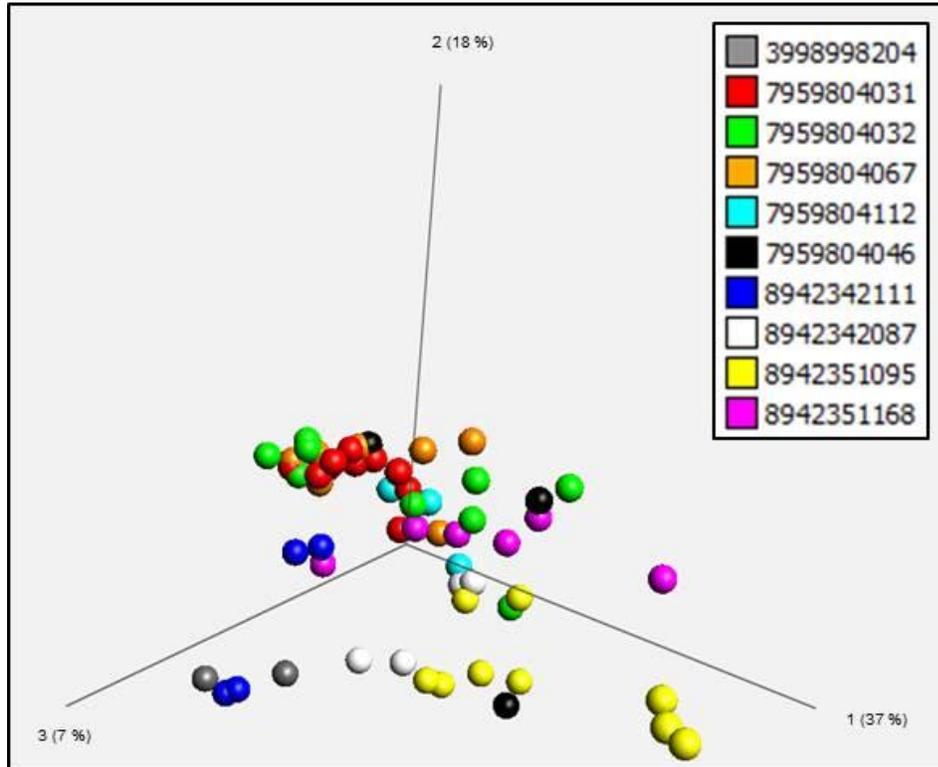


Abbildung 60 Unsupervidierte Analyse als PCA-Darstellung (Varianzfilter $\sigma/\sigma_{\max} \geq 0,4$) der AVG-Beta-Werte von 60 Herzgewebeproben verschiedener kardialer Regionen, die mittels 450K-Array analysiert wurden. Die 450K-Array-ID wurde zur Analyse von *Batch*-Effekten untersucht. Dargestellt ist die Segregation der 60 Herzgewebeproben basierend auf 7763/447821 CpG-Loci mit $\sigma/\sigma_{\max} \geq 0,4$. Der Farbcode ist nach 450K-Array-ID dargestellt. Mithilfe der PCA-Darstellung sollte überprüft werden, ob Proben, die auf demselben Array hybridisiert wurden, zusammen gruppieren und somit Batch-Effekte zeigen. Hier ist keine Gruppierung bzw. Segregation entsprechend der 450K-Array-ID zu erkennen, sodass keine Batch-Effekte vorlagen.

Gen		IRX4		FGFR2	
CpG-Locus		cg17774559		cg17710779	
CHR-Position		chr5:1879698		chr10:123356331	
Probe Gruppe	Probe Bezeichnung	450K	Epityper	450K	Epityper
Atrium	HLHS-TGA_LA1	0.20	0.14	0.18	0.29
	HLHS-TGA_RA1	0.28	0.09	0.33	0.21
	Lvan_LA1	0.11	0.09	0.31	0.29
	Lvan_RA1	0.14	0.04	0.15	0.14
	HLHS_4_IAS	0.29	0.22	0.14	0.13
	HLHS_182_IAS		0.27	0.19	0.08
Ventrikel	HLHS-TGA_LV1	0.53			NA
	HLHS-TGA_RV1	0.64			
	Lvan_LV1	0.70	0.54	0.52	0.53
	Lvan_RV1	0.79	0.59	0.64	0.60
Methylierte Kontrolle		-	0.98	-	0.87
Unmethylierte Kontrolle		-	0.08	-	0.08
Differenz Atrium - Ventrikel		-0.40	-0.37	-0.31	-0.32

Abbildung 61 Gegenüberstellung der DNA-Methylierungswert-Rohdaten der DM-CpGs der Gene *IRX4* (cg17774559) und *FGFR2* (cg17710779) von 450K-Array und der Epityper-Analyse. Die aufgeführten CpG-Loci lagen signifikant differenziell methyliert zwischen atrialen und ventrikulären Herzgewebetypen in der 450K-Analyse vor. Zur Verifizierung mittels Epityper-Technik wurden 6 DNA-Proben des „Atrium“-Kollektivs und vier DNA-Proben des „Ventrikel“-Kollektivs analysiert. Eine methylierte und eine unmethylierte Kontrolle (EpigenDx, siehe 2.2.1) wurden für die Epityper-Technik verwendet und ergaben entsprechend einen hohen Methylierungswert von ~ 0,9 bei methylierter Kontrolle und einen minimalen Methylierungswert von ~ 0,1 bei unmethylierter Kontrolle. Insgesamt fielen hierbei die bestimmten Methylierungswerte bei der Epityper-Technik geringer aus, als die 450K-Methylierungswerte. Die DNA-Methylierungswerte in „Ventrikel“ bei *IRX4* (cg17774559) konnten durch die Epityper-Technik teilweise nur bedingt nachgewiesen werden. Die Differenz der DNA-Methylierung (Abbildung 36: „Differenz Atrium – Ventrikel“) zwischen dem Mittelwert der DNA-Methylierungswerte der atrialen und ventrikulären Herzgewebeproben blieb jedoch tendenziell gleich. Differenz Atrium - Ventrikel: Differenz der DNA-Methylierung zwischen dem Mittelwert der DNA-Methylierungswerte der „Atrium“-Proben und der der „Ventrikel“-Proben. NA: kein Messwert vorhanden. Farbcode DNA-Methylierung: blau (unmethyliert) bis gelb (voll methyliert).

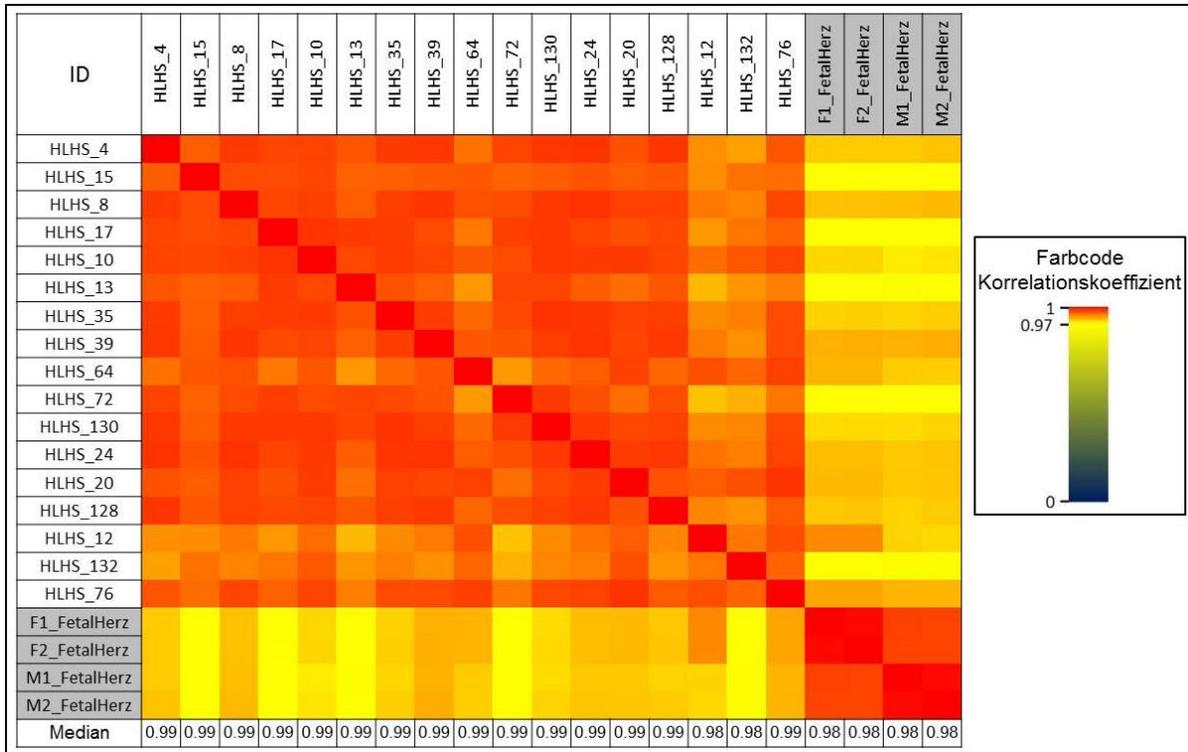


Abbildung 62 CpG-Gesamtkorrelation der AVG-Beta-Werte von 17 IAS-Gewebe von HLHS-Patienten und Duplikaten von fetalem Herzgewebe je eines männlichen (M1/M2) und eines weiblichen (F1/F2) Donors, die auf einem 450K-Array hybridisiert wurden. Dargestellt sind die Korrelationskoeffizient-Werte (r) mit einer farblich bedingten Formatierung von niedrigen r-Werten in blau und hohen r-Werten in rot (siehe Farbcode). Jede Spalte (und die gleiche Zeilennummer) steht für eine DNA-Probe, mit oberhalb und links der Korrelationsmatrix den Herzgewebeproben HLHS in weiß und fetales Herzgewebe in grau. Unterhalb der Korrelationsmatrix ist der Median über alle Korrelationskoeffizienten der jeweiligen Spalte/DNA-Probe dargestellt.

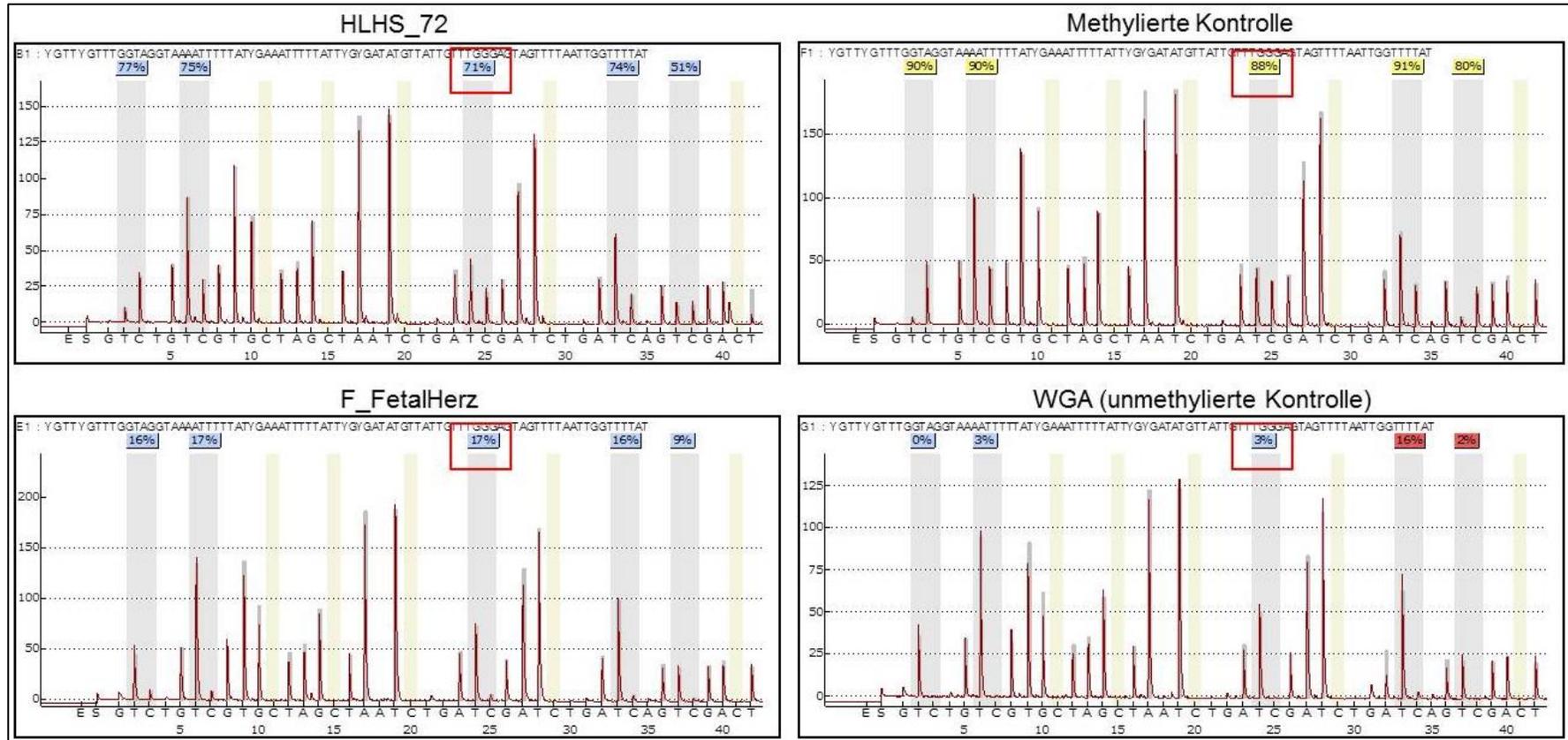


Abbildung 63 Pyrogramm-Darstellung (PyroQ-CpGTM Software) des CpG-Locus cg04192879 von *NR2F1* der Proben HLHS_72, F_FetalHerz, methylierte Kontrolle (hochmethylierte (> 85 %) Kontroll-DNA (EpigenDx)) und WGA (unmethylierte (< 5 %), WGA-amplifizierte DNA). Oben ist die zu analysierende bisulfit-konvertierte Sequenz dargestellt, x-Achse: Dispersion Order der eingespritzten dTNPs, sowie Enzym-(E) und Substratmix (S), y-Achse: Signalintensität; rote Peaks: detektierte Signalintensitäten; Einbau zweier gleicher Nucleotide ergibt eine doppelte peak-Höhe; graue Balken hinter roten Peaks: Histogramm der theoretischen Signalintensitäten; oranger Balken: Konversionskontrolle (kein Peak erwartet), graue breite Balken: CpG-Position, Prozent-Angabe: DNA-Methylierung [%] am entsprechenden CpG-Locus. Der zu analysierende CpG-Locus ist rot eingerahmt. Zusätzliche Informationen zur Pyrogramm-Darstellungen sind in Abschnitt 2.2.2.3 aufgeführt.

Exom						Epigenom				
Fall	Gen	Mutation Position CHR (GRCh37/hg19)	Typ	PPH	ExAC AF	CpG ID Position CHR (GRCh37/hg19)	DHS	Enhancer- Assoziation	DNA-Methylierung [AVG-Beta]	
									Muskel	Gefäß
HLHS_35	ANTXR1	c.1508C>A chr2:69472430	Miss.	B	3.02E-05	cg24576133 chr2:69299868	-	ja		
						cg24676755 chr2:69346341	-	ja		
HLHS_132	CLIP2	c.1042G>A chr7:73771634	Miss.	D	-	cg11573518 chr7:73721031	ja	ja		
HLHS_13	COL17A1	c.875C>T chr10:105822051	Miss.	-	-	cg04215238 chr10:105841897	-	ja		
HLHS_21/22	CYB5R3	c.260G>A chr22:43027449	Miss.	-	-	cg21827203 chr22:43016886	ja	-		
HLHS_132	DOCK1	c.5038A>G chr10:129237331 c.1927G>A chr10:128836060	Miss.	D	-	cg15676916 chr10:129191358	-	-		
			Miss.	P	-					
HLHS_12	DYRK2	c.1120T>C chr12:68051807	Miss.	D	-	cg09136343 chr12:68041428	-	-		
HLHS_15	FAM53B	c.461G>A chr10:126370621	Miss.	P	4.61E-05	cg17879101 chr10:126329354	-	ja		
HLHS_8	FAM65A	c.1200_1206delCAGCCCA chr16:67575847-67575853	FS	-	-	cg05362168 chr16:67562128	-	-		
HLHS_35	ITPR1	c.3412A>G chr3:4725441	Miss.	-	0.00085	cg02105211 chr3:4625188	-	ja		
HLHS_8	LAMB4	c.3601G>C chr7:107696231	Miss.	B	-	cg26395761 chr7:107700124	-	ja		
HLHS_4	MGMT	c.611G>T chr10:131565155	Miss.	-	6.16E-05	cg13171843 chr10:131324189	ja	-		
HLHS_21/22	OAT	c.718G>A chr10:126092420 c.1517C>G chr5:157065601	Miss.	B	1.50E-05	cg04275847 chr10:126105697	-	-		
HLHS_8	SOX30	c.1369C>T chr2:107429630	Miss.	D	-	cg09842125 chr5:157080710	-	-		
HLHS_13	ST6GAL2	c.385T>G chr10:134041566	Miss.	D	-	cg08775629 chr2:107457756	-	-		
HLHS_8	STK32C	c.385T>G chr10:134041566	Miss.	D	-	cg03435248 chr10:134079425	ja	ja		
HLHS_132	TNS3	c.776G>A chr7:47440459	Miss.	D	3.00E-05	cg23026754 chr7:47579678	-	-		
						cg27275374 chr7:47539765	-	-		
						cg23895495 chr7:47539837	-	-		
						cg04879211 chr7:47576144	ja	ja		

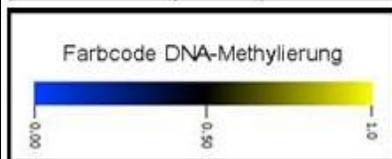


Abbildung 64 Mutation und DM-CpG-Loci zu 16 überlappenden Genen aus der WES-Analyse von IAS-Gewebe von neun HLHS-Patienten (siehe 3.2.1.1) und DNA-Methylierungsanalyse „Muskel-Gefäß“ (siehe 3.2.2.4) eines DCM-Patienten. Zu jedem HLHS-Fall sind Genvarianten und überlappende DM-CpG-Loci mit Darstellung der nach Farbcode-Legende eingefärbten AVG-Beta-Mittelwerte der Gewebegruppen „Muskel“ (LA, RA, IAS, LV, RV, TV, MV) vs. „Gefäß“ (PA, Aorta, AoV). Miss.: Missense-Mutation, FS: Frameshift-Mutation. PPH: prädiktive Mutationskonsequenz nach Polyphen-2. B = *benign*, P = *possibly damaging*, D = *probably damaging*. ExAC AF: Allelfrequenz nach ExAC (Version 0.3). E= Zehnerpotenz. DHS (*DNase I hypersensitive site*) und Enhancer-Assoziation nach Illumina Annotation.

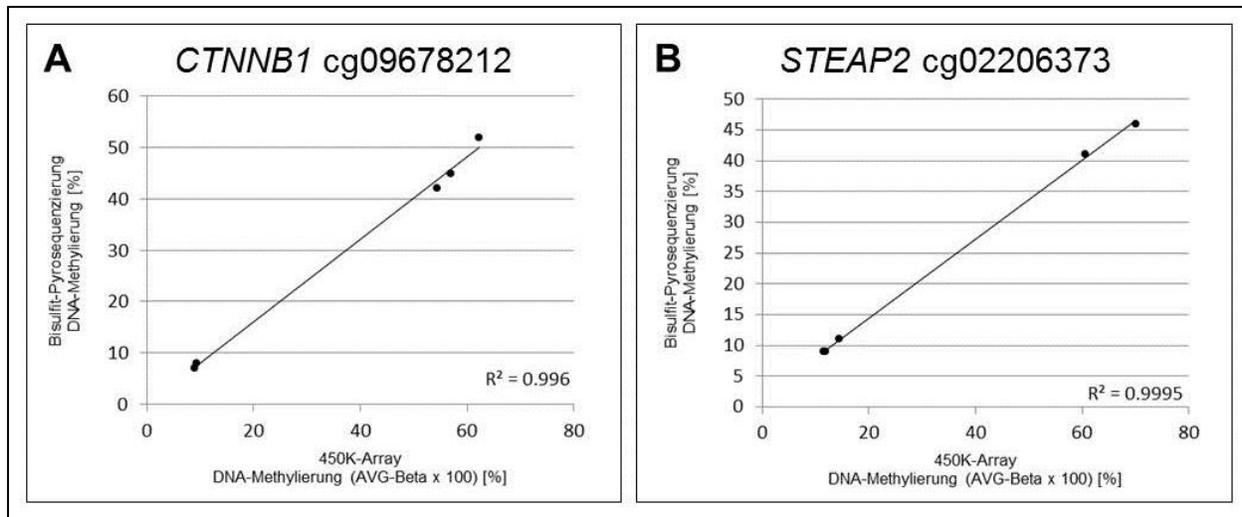


Abbildung 65 Lineare Regression der DM-CpGs der Gene (A) *CTNNB1* (cg09678212) und (B) *STEAP2* (cg02206373) nach Korrelationsanalyse (Pearson-Methode) der DNA-Methylierungswerte des 450K-Arrays und der BSPS. Die aufgeführten CpG-Loci lagen differenziell methyliert zwischen den Gruppen HLHS und fetalem Herzkontrollgewebe in der 450K-Analyse vor. Zur Verifizierung mittels BSPS wurden drei DNA-Proben des HLHS-Kollektivs (HLHS_13, HLHS_72 und HLHS_130) und die beiden DNA-Proben der fetalen Herzgewebe (M_FetalHerz und F_FetalHerz) analysiert. Methylierungswerte der BSPS (DNA-Methylierung [%]) auf der y-Achse wurden gegen die 450K-Methylierungswerte (DNA-Methylierung (AVG-Beta x 100) [%]) auf der x-Achse in einem Scatterplot als lineare Regression dargestellt. Das Bestimmtheitsmaß R^2 ergab für beide analysierte CpG-Loci $R^2 > 0,99$.

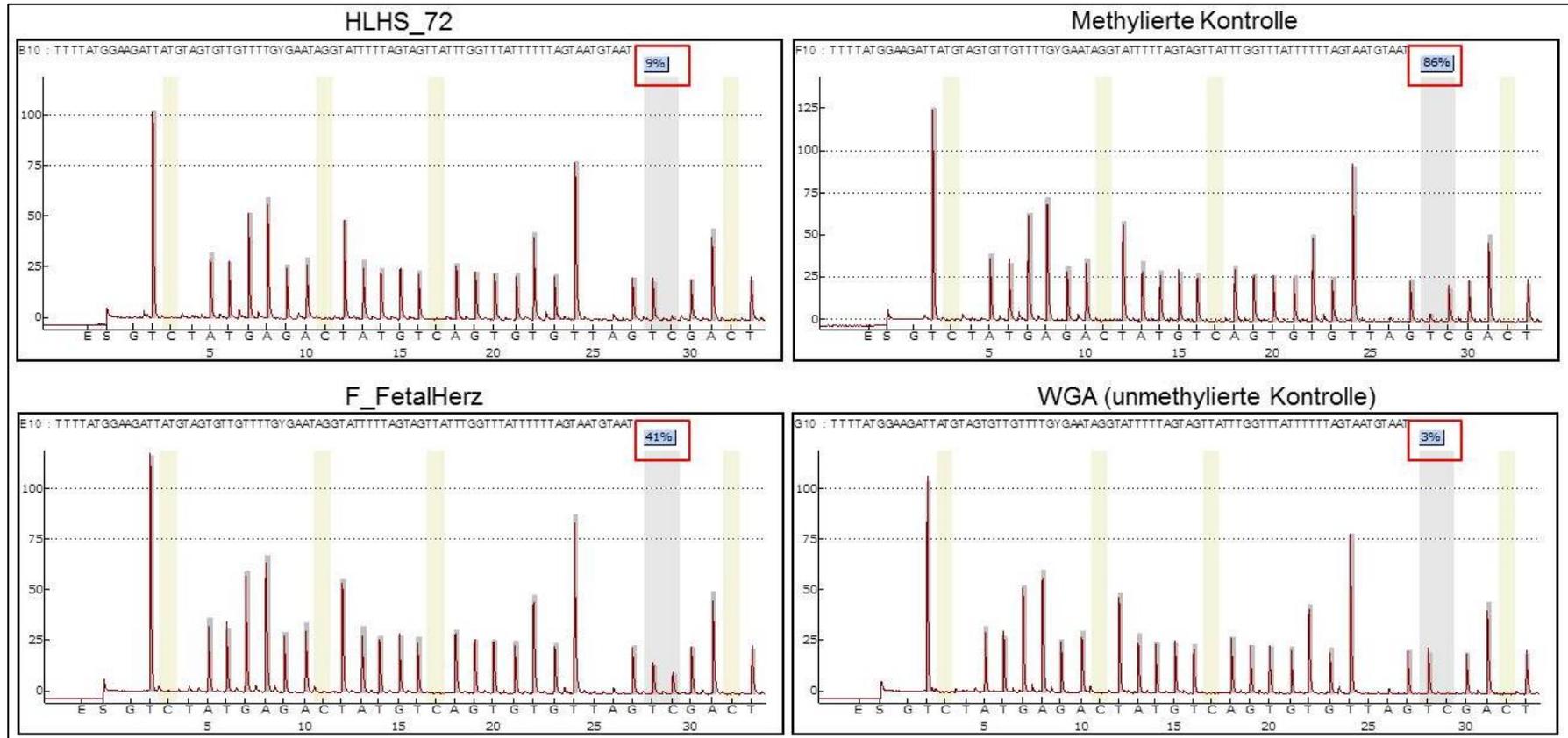


Abbildung 66 Pyrogramm-Darstellung (PyroQ-CpG™ Software) des CpG-Locus cg02206373 von *STEAP2* der Proben HLHS_72, F_FetalHerz, methylierte Kontrolle (hochmethylierte (> 85 %) Kontroll-DNA (EpigenDx)) und WGA (unmethylierte (< 5 %), WGA-amplifizierte DNA). Oben ist die zu analysierende bisulfite-konvertierte Sequenz dargestellt, x-Achse: Dispensation Order der eingespritzten dNTPs, sowie Enzym-(E) und Substratmix (S), y-Achse: Signalintensität; rote Peaks: detektierte Signalintensitäten; Einbau zweier gleicher Nukleotide ergibt eine doppelte peak-Höhe; graue Balken hinter roten Peaks: Histogramm der theoretischen Signalintensität, oranger Balken: Konversionskontrolle (kein Peak erwartet), graue breite Balken: CpG-Position, Prozent-Angabe: DNA-Methylierung [%] am entsprechenden CpG-Locus. Der zu analysierende CpG-Locus ist rot eingrahmt. Zusätzliche Informationen zur Pyrogramm-Darstellungen sind in Abschnitt 2.2.2.3 aufgeführt.

9 Danksagung

Ich möchte mich herzlich bei allen bedanken, die mich bei der Anfertigung meiner Dissertation auf vielfältige Weise unterstützt haben:

Bei Herrn Prof. Dr. rer. nat. Hinrich Schulenburg für die Übernahme des Hauptreferates,

bei Herrn Prof. Dr. med. Hans-Heiner Kramer für die Betreuung meiner Doktorarbeit und die kritische Durchsicht dieser Arbeit. Weiterhin bedanke ich mich für das interessante Dissertationsthema und die sich daraus ergebene Mitwirkung an derzeitigen und zukünftigen spannenden Projekten zur Thematik der kongenitalen Herzfehlbildungen.

Des Weiteren möchte ich mich bei Herrn Prof. Dr. med. Reiner Siebert und bei Herrn Prof. Dr. rer. nat. Ole Ammerpohl für die Unterstützung und Betreuung während meiner Doktorarbeit bedanken.

Ich möchte allen Kooperationspartnern des Kompetenznetz Angeborene Herzfehler e.V. für die Zusammenarbeit danken. Insbesondere bedanke ich mich bei Dr. rer. nat. Thomas Pickardt, der mich beim Akquirieren des Probenmaterials tatkräftig unterstützt hat.

Mein besonderer Dank gilt Herrn Dr. med. Marc-Phillip Hitz und seiner facettenreichen Unterstützung. Seine fachliche Expertise und organisatorische Unterstützung ermöglichten die Partizipation an neuen interessanten Projekten. Weiterhin möchte ich mich bei Frau Dr. med. Anne-Karin Kahlert für die konstruktive Zusammenarbeit und hilfreichen Ratschläge bedanken.

Für die gute Zusammenarbeit, das harmonische Arbeitsumfeld und die Hilfe bei der Durchführung experimenteller Arbeiten möchte ich mich herzlich bei Claudia Becher und Lorena Valles bedanken.

Insbesondere möchte ich mich bei Kerstin Runde bedanken. In Zusammenarbeit mit ihr ist es möglich gewesen, den Aufbau einer Zebrafisch-Laboreinheit in der Klinik für angeborene Herzfehler und Kinderkardiologie zu meistern.

Mein allergrößter Dank gilt meinem Freund Marius Wolfgramm, der mir immer mit Rat und Zuspruch zur Seite steht und mir viel Kraft und Unterstützung in allen Phasen dieser Doktorarbeit gegeben hat.

10 Lebenslauf

Persönliche Daten

Name: Kirstin Hoff
 Akademischer Grad: Master of Science (M. Sc.) Biomedizin
 Anschrift: Eckernförder Str. 64
 24116 Kiel
 Geburtstag: 23.12.1985
 Geburtsort: Hannover
 Familienstand: ledig

Schulbildung und beruflicher Werdegang

1992 – 1996 Grundschule, Hannover
 1996 – 1998 Orientierungsstufe, Hannover
 1998 – 2005 Gymnasium, Hannover
 08/2002 – 12/2002 Auslandsaufenthalt: erstes Halbjahr der Vorstufe an der „*Strathcona-Tweedsmuir School*“ in Calgary, Kanada
 2005 Abschluss: Allgemeine Hochschulreife mit naturwissenschaftlicher Ausrichtung
 10/2005 – 02/2006 Studiengang Medizin, ein Semester an der Universität Hamburg
 02/2006 – 04/2006 Betriebstechnisches Praktikum
 10/2006 – 06/2009 Studiengang Pflanzen-Biotechnologie an der Leibniz Universität Hannover mit Abschluss Bachelor of Science
 08/2009 Auszeichnung durch Albert-Ludwig-Fraas-Stiftung
 10/2009 – 09/2011 Studiengang Biomedizin an der Medizinischen Hochschule Hannover mit Abschluss Master of Science
 Vertiefungsrichtung: Immunologie und Mikrobiologie
 03/2011 – 09/2011 Masterarbeit am Institut für Mikrobiologie der Medizinischen Hochschule Hannover:
 „*Reinigung und Charakterisierung des putativen Helicobacter pylori cagPAI-Effektorproteins CagN und Optimierung der Reinigung des cagPAI-Strukturproteins CagL*“
 04.11.2011 Auszeichnung mit dem Studienpreis im Masterstudiengang Biomedizin
 seit 01/2012 wissenschaftliche Angestellte an der Klinik für angeborene Herzfehler und Kinderkardiologie, Universitätsklinikum Schleswig-Holstein, mit Zielrichtung Promotion
 02/2016 Absolvierung der Ausbildung zum Projektleiter gentechnischer Arbeiten gemäß § 15 Gentechnik-Sicherheitsverordnung

11 Publikationsliste

(*: geteilte Erstautorenschaft)

Publizierte Originalarbeiten:

Ehmke N.*, Caliebe A.*, Koenig R., Kant S.G., Stark Z., Cormier-Daire V., Wieczorek D., Gillissen-Kaesbach G., **Hoff K.**, Kawalia A., Thiele H., Altmüller J., Fischer-Zirnsak B., Knaus A., Zhu N., Heinrich V., Huber C., Harabula I., Spielmann M., Horn D., Kornak U., Hecht J., Krawitz P. M., Nürnberg P., Siebert R., Manzke H., Mundlos S. (2014). Homozygous and Compound-Heterozygous Mutations in TGDS Cause Catel-Manzke Syndrome. *Am. J. Hum. Genet.* 95, 763–770.

Al Turki S. H.*, Manickaraj A. K.*, Mercer C. L.*, Gerety S. S.*, Hitz M.-P., Lindsay S., D'Alessandro L. C. A., Swaminathan G. J., Bentham J., Arndt A.-K., Low J., Breckpot J., Gewillig M., Thienpont B., Abdul-Khaliq H., Harnack C., **Hoff K.**, Kramer H.-H., Schubert S., Siebert R., Toka O., Cosgrove C., Watkins H., Lucassen A. M., O'Kelly I. M., Salmon A. P., Bu'Lock F. A., Granados-Riveron J., Setchfield K., Thornborough C., Brook J. D., Mulder B., Klaassen S., Bhattacharya S., Devriendt K., Fitzpatrick D. R., UK10K Consortium, Wilson D. I., Mital S.*, and Hurles M. E.* (2014). Rare Variants in NR2F2 Cause Congenital Heart Defects in Humans. *Am. J. Hum. Genet.* 94, 574–585.

Arbeiten, die in Revision sind:

Sifrim A.*, Hitz M.-P.*, Wilsdon A., Breckpot J., Al Turki S. H., Thienpont B., McRae J., Fitzgerald T. W., Singh T., Swaminathan G. J., Prigmore E., Rajan D., Abdul-Khaliq H., Banka S., Bauer U. M. M., Bentham J., Berger F., Bhattacharya S., Bu'Lock F., Canham N., Colgiu I.-G., Cosgrove C., Cox H., Daehner I., Daly A., Danesh J., Fryer A., Gewillig M., Hobson E., **Hoff K.**, Homfray T., The INTERVAL Study, Kahlert A.-K., Ketley A., Kramer H.-H., Lachlan K., Lampe A. K., Louw J. J., Manickara A. K., Manase D., McCarthy K. P., Metcalfe K., Moore C., Newbury-Ecob R., Omer S. O., Ouwehand W. H., Park S.-M., Parker M. J., Pickardt T., Pollard M. O., Robert L., Roberts D. J., Sambrook J., Setchfield K., Stiller B., Thornborough C., Toka O., Watkins H., Williams D., Wright M., Mital S., Daubeney P., Keavney B., Goodship J.,

The UK10K Consortium, Abu-Sulaiman R. M., Klaassen S., Wright C. F., Firth H. V., Barrett J. C., Devriendt K., Fitzpatrick D. R., Brook J. D., The Deciphering Developmental Disorders Study, Hurles M. E. (2016). Exome sequencing reveals distinct genetic architectures for syndromic and nonsyndromic congenital heart defects. *Nature Genetics*, under revision

Bereits publizierte Arbeiten, die nicht Bestandteil der Dissertation sind:

Seth S., Oberdörfer L., Hyde R., Hoff K., Thies V., Worbs T., Schmitz S., Förster R. (2011). CCR7 essentially contributes to the homing of plasmacytoid dendritic cells to lymph nodes under steady-state as well as inflammatory conditions. *J Immunol.* 15;186(6):3364-72.

Bens S., Zichner T., Stütz A.M., Caliebe A., Wagener R., Hoff K., Korbel J.O., von Bismarck P., Siebert R (2014). SPAG7 is a candidate gene for the periodic fever, aphthous stomatitis, pharyngitis, and adenopathy (PFAPA) syndrome. *Genes Immun.* 15(3):190-4.

Kürzlich eingereichte Arbeiten:

Kahlert A.-K., Hoff K., Hitz M.-P. (2016). Genetik der angeborenen strukturellen Herzfehler. *Medizinische Genetik*, eingereicht am 09.03.2016

Arbeiten, die in Vorbereitung für eine zukünftige Publikation sind:

Hitz M.-P., Hoff K., Al Turki S., Sifrim S., Schalinski A., Bauer U., Pickardt T., Hurles M., Klaassen S. (Autorenkombination ändert sich noch) Thematik: Whole Exome Sequencing (WES) in Families with Left Ventricular Outflow Tract Obstruction (LVOTO).

Hoff K. „et al.“ (Autoren stehen noch nicht fest) Thematik: Differential DNA Methylation Patterns in Cardiac Tissue from Atrium and Ventricles.

Publizierte Tagungsbeiträge:

Hoff K., Ammerpohl O., Kolarova J., Arndt A.-K., Pfeffer K., OToka O., Siebert R., Kramer H.-H. (2013). Determination of DNA Methylation Patterns in Cardiac Tissue from Hypoplastic Left Heart Syndrome. 45. Jahrestagung der Deutschen Gesellschaft für Pädiatrische Kardiologie (DGPK), 2. – 5. November 2013, Weimar, Deutschland. *Thorac cardiovasc Surg.* 61 - P10

Arndt A.-K., Musso G., Gellert P., **Hoff K.**, Uchida S., Toka O., Siebert R., MacRae C. A., Klaassen S., Kramer H.-H. (2013). Transcriptome sequencing in hypoplastic left heart syndrome myocardium reveals dysregulated cell cycle genes. 45. Jahrestagung der Deutschen Gesellschaft für Pädiatrische Kardiologie (DGPK), 2. – 5. November 2013, Weimar, Deutschland. *Thorac cardiovasc Surg.* 61 - V16

Hoff, K. Ammerpohl O., Kolarova J., Arndt A.-K., Thiele H., Pfeffer K., Toka O., Nürnberg P., Siebert R., Kramer H.-H. (2014). Exome Sequencing and DNA Methylation Analysis of Cardiac Tissue from Patients with Hypoplastic Left Heart Syndrome. Vortrag bei der 25. Jahrestagung der Deutschen Gesellschaft für Humangenetik (GfH), 19. – 21. März 2014, Essen, Deutschland. *Medizinische Genetik*, Band 25, Heft 1, 2014.

12 Erklärung

Hiermit erkläre ich, dass die vorliegende Dissertation, abgesehen von der Beratung durch betreuende Personen, nach Inhalt und Form das Resultat meiner eigenen Arbeit ist. Diese Dissertation wurde bis jetzt weder ganz noch zum Teil einer anderen Stelle im Rahmen eines Promotionsverfahrens vorgelegt.

Einige Teilergebnisse dieser Dissertation wurden bereits publiziert bzw. sind derzeit in Bearbeitung für eine zukünftige Publikation (siehe Abschnitt 11 Publikationsliste).

Die vorliegende Arbeit ist unter Einhaltung der Regeln guter wissenschaftlicher Praxis der Deutschen Forschungsgemeinschaft entstanden.

Kiel, den