



# **Genetische Analyse kardiovaskulärer Erkrankungen**



DISSERTATION ZUR ERLANGUNG DES DOKTORGRADES DER  
NATURWISSENSCHAFTEN (DR. RER. NAT.) DER FAKULTÄT FÜR BIOLOGIE  
UND VORKLINISCHE MEDIZIN DER UNIVERSITÄT REGENSBURG

vorgelegt von

Ulrike Esslinger

aus

Pfarrkirchen

im Jahr 2012

Das Promotionsgesuch wurde eingereicht am: 17.04.2012

Die Arbeit wurde angeleitet von: Prof. Dr. Christian Hengstenberg

Unterschrift:

**Diese Arbeit ist meinem Vater  
gewidmet.**

# Inhaltsverzeichnis

Inhaltsverzeichnis .....	I
Abbildungsverzeichnis .....	IV
Tabellenverzeichnis .....	VI
Spezielle Abkürzungen der Arbeit .....	VIII
1. Einleitung .....	1
1.1. Epidemiologie kardiovaskulärer Erkrankungen .....	1
1.2. Atherosklerose und Herzinfarkt .....	2
1.3. Kardiomyopathien .....	10
1.3.1. Dilatative Kardiomyopathie .....	11
1.3.2. Hypertrophische Kardiomyopathie .....	13
1.4. Geschichte der Sequenzierung .....	16
1.5. Zielsetzung .....	19
2. Material .....	20
2.1. Geräte .....	20
2.2. Verbrauchsmaterialien .....	21
2.3. Software .....	22
3. Methoden .....	23
3.1. Rekrutierung .....	23
3.1.1. Rekrutierungsstrategie Herzinfarkt-Großfamilien .....	23
3.1.2. Rekrutierungsstrategie EUROGENE- <i>Heart Failure Project</i> .....	23
3.1.3. Auswahl der Familien für diese Arbeit .....	24
3.2. Molekularbiologische Methoden .....	24
3.2.1. DNA-Isolierung und Konzentrationsbestimmung .....	24
3.2.2. Oligonukleotide .....	24
3.2.3. Amplifikation von DNA-Fragmenten mittels PCR .....	25
3.2.4. Touchdown-PCR .....	26
3.2.5. Agarosegelelektrophorese .....	26
3.2.6. Isolierung und Aufreinigung von DNA-Fragmenten .....	27
3.2.7. DNA-Sequenzierung .....	27
3.2.8. Gelelektrophorese mit Bioanalyser .....	27
3.3. Exom-Sequenzierung .....	28

---

3.4.	Erstellung der Exom-Bibliotheken.....	29
3.4.1.	Erstellung der Exom-Bibliothek nach Agilent Technologies .....	29
3.4.2.	Erstellung der Exom-Bibliothek nach Illumina und Vergleich .....	31
3.4.3.	Modifiziertes Protokoll zur Erstellung der Exom-Bibliothek nach Illumina .....	32
3.5.	Hochdurchsatz-Sequenzierungsplattform Illumina.....	33
3.5.1.	Cluster Erzeugung für Sequenzierung.....	33
3.5.2.	Sequenzierungsplattform Illumina .....	34
3.5.3.	Cluster Erzeugung und Sequenzierung Genome Analyzer Iix .....	35
3.5.4.	Cluster Erzeugung und Sequenzierung HiSeq2000 .....	35
3.6.	Auswertung der Exom-Sequenzen.....	36
3.6.1.	Erwartungsbereich der Information für generierte Exom-Sequenzen.....	36
3.6.2.	Visualisierung mit dem Integrative Genomics Viewer (IGV).....	37
3.6.3.	Bioinformatische Auswertungsstrategie 1 und 2 .....	38
3.6.4.	Zusätzliche Qualitätsfilter für beide Strategien.....	39
3.6.5.	Filtern von paralogen Genen .....	40
3.6.6.	Vorhersagemodul PolyPhen-2.....	41
3.7.	Mutationsscreening, Genotypisierung und Expression .....	41
3.7.1.	IDAHO-Screening .....	41
3.7.2.	Genotypisierung mit KASPar Assay .....	44
3.7.3.	Quantitative PCR .....	45
3.7.4.	Berechnung der relativen Genexpression.....	46
4.	Ergebnisse.....	47
4.1.	Exom-Sequenzierung Herzinfarkt-Großfamilie .....	47
4.1.1.	Exom-Sequenzierung HI-Großfamilie 2001 .....	47
4.1.2.	Identifizierte Varianten.....	49
4.1.3.	Ko-Segregation der identifizierten Varianten .....	50
4.1.4.	Einblick und Daten der CARDIoGRAM-Studie.....	51
4.1.5.	IDAHO-Screening für ein Kandidatengen .....	53
4.1.6.	Sequenzierung weiterer Kandidatengene.....	54
4.2.	Fragmentierung der DNA für Exom-Sequenzierung DCM- und HCM-Familien.....	56
4.3.	Exom-Sequenzierung DCM-Familie 9800 .....	57
4.4.	Exom-Sequenzierung DCM-Familie 16725 .....	62
4.4.1.	Identifizierte Varianten.....	64
4.4.2.	Ko-Segregation der <i>TTN</i> -Variante .....	67
4.5.	Exom-Sequenzierung DCM-Familie 009-1229 .....	70
4.5.1.	Identifizierte Varianten der modifizierten und nicht modifizierten Proben.....	71

---

4.5.2.	Zusätzlich identifizierte Varianten der nicht modifizierten Proben .....	79
4.5.3.	Zusätzlich identifizierte Varianten der modifizierten Proben.....	82
4.6.	Exom-Sequenzierung HCM-Familie 757 .....	85
4.6.1.	Identifizierte Varianten.....	87
4.6.2.	Ko-Segregation der <i>HOOX1</i> -Variante und Expression .....	89
4.7.	Gesamtauswertung der DCM- und HCM-Familien.....	91
5.	Diskussion .....	93
5.1.	Fehlende Heritabilität kardiovaskulärer Erkrankungen.....	93
5.2.	Exom-Sequenzierung zur Identifizierung seltener Varianten .....	95
5.3.	Exom-Sequenzierung in einer HI-Großfamilie .....	100
5.4.	Exom-Sequenzierung in DCM- und HCM-Familien.....	105
5.5.	Ausblick.....	109
6.	Zusammenfassung .....	111
7.	Literaturverzeichnis.....	112
8.	Anhang .....	127
8.1.	Klinische Charakteristika der Familien dieser Arbeit .....	127
8.2.	Klinische Charakteristika der Screening-Kollektive .....	132
8.3.	Oligonukleotide .....	133
8.4.	Parameter von SAMtools.....	135
8.5.	Stammbaum HI-Großfamilie 2001 .....	136
8.6.	Anzahl detektierter Varianten .....	137
	Danksagung .....	143
	Eidesstaatliche Erklärung .....	144

## Abbildungsverzeichnis

<b>Abbildung 1:</b> Bildung von <i>fatty streaks</i> in Atherosklerose .....	2
<b>Abbildung 2:</b> Ruptur eines Plaques.....	3
<b>Abbildung 3:</b> Risikoerhöhung für KHK/HI in Anbetracht der familiären Disposition...	4
<b>Abbildung 4:</b> Manhattan-Plot der Ergebnisse der CARDIoGRAM-Metaanalyse .....	8
<b>Abbildung 5:</b> Schematische Darstellung eines Sarkomers.....	10
<b>Abbildung 6:</b> Darstellung von einem gesunden Herzen/Herzen mit DCM.....	11
<b>Abbildung 7:</b> Darstellung von einem gesunden Herzen/Herzen mit HCM.....	13
<b>Abbildung 8:</b> Proteine, in denen kausale Mutationen für HCM gefunden wurden ...	14
<b>Abbildung 9:</b> Erstellung einer Bibliothek für die Exom-Sequenzierung – Teil 1.....	29
<b>Abbildung 10:</b> Adapter für <i>paired-end</i> Sequenzierung mit und ohne Index.....	29
<b>Abbildung 11:</b> Erstellung einer Bibliothek für die Exom-Sequenzierung – Teil 2.....	30
<b>Abbildung 12:</b> cBot und <i>flowcell</i> .....	33
<b>Abbildung 13:</b> Erzeugung von Cluster.....	33
<b>Abbildung 14:</b> Sequenzierung durch Synthese. ....	34
<b>Abbildung 15:</b> Visualisierung der Exom-Sequenzen mit IGV .....	37
<b>Abbildung 16:</b> Auswertung der Schmelzkurven eines IDAHO-Screenings.....	43
<b>Abbildung 17:</b> Stammbaum HI-Großfamilie 2001 .....	47
<b>Abbildung 18:</b> Konservierung der Mutationsstelle der <i>GUCY1A3</i> -Insertion .....	50
<b>Abbildung 19:</b> Assoziationsplots aus der CARDIoGRAM-Studie .....	52
<b>Abbildung 20:</b> Bioanalyser-Profile fragmentierter DNA .....	56
<b>Abbildung 21:</b> Stammbaum der DCM-Familie 9800.....	57
<b>Abbildung 22:</b> Anzahl ausgewählter Varianten der DCM-Familie 9800.....	59
<b>Abbildung 23:</b> Stammbaum der DCM-Familie 16725.....	62
<b>Abbildung 24:</b> Anzahl ausgewählter Varianten der DCM-Familie 16725.....	65
<b>Abbildung 25:</b> Genotypisierung der <i>TTN</i> -Variante in der DCM-Familie 16725.....	67
<b>Abbildung 26:</b> Konservierung der Mutationsstelle der <i>TTN</i> -Variante .....	69
<b>Abbildung 27:</b> Stammbaum der DCM-Familie 009-1229.....	70
<b>Abbildung 28:</b> Anzahl ausgewählter Varianten der DCM-Familie 009-1229 .....	71
<b>Abbildung 29:</b> Konservierung der Mutationsstelle der <i>WHAMM</i> -Variante.....	78
<b>Abbildung 30:</b> Relative Expression von <i>GJB4</i> .....	78
<b>Abbildung 31:</b> Relative Expression von <i>ZG16</i> und <i>INTS1</i> .....	79
<b>Abbildung 32:</b> Stammbaum der HCM-Familie 757.....	85



---

<b>Abbildung 33:</b> Konservierung der Mutationsstelle der <i>HOOK1</i> -Variante.....	90
<b>Abbildung 34:</b> Relative Expression von <i>HOOK1</i> . .....	90
<b>Abbildung 35:</b> Allelfrequenz und Effektgröße identifizierter Varianten .....	94
<b>Abbildung 36:</b> <i>Alignment</i> -Qualität einer Variante im Gen <i>PGM2</i> .....	97
<b>Abbildung 37:</b> Ausschnitt aus dem Stammbaum der HI-Großfamilie 2001 .....	101
<b>Abbildung 38:</b> NO-sGC-cGMP Signalweg in einem Blutgefäß.....	102
<b>Abbildung 39:</b> Komplexität kardiovaskulärer Phänotypen.....	110
<b>Abbildung 40:</b> Stammbaum HI-Großfamilie 2001 .....	136

---

## Tabellenverzeichnis

<b>Tabelle 1:</b> GWAS von häufigen kardiovaskulären Erkrankungen.....	7
<b>Tabelle 2:</b> Ontologie kausaler DCM-Gene .....	12
<b>Tabelle 3:</b> Kausale Gene für HCM .....	15
<b>Tabelle 4:</b> Vergleich der Erstellung der Exom-Bibliotheken .....	31
<b>Tabelle 5:</b> Vergleich der bioinformatischen Auswertung .....	38
<b>Tabelle 6:</b> Annotationskategorien bei ANNOVAR .....	39
<b>Tabelle 7:</b> Vergleich der zusätzlichen Qualitätsfilter.....	40
<b>Tabelle 8:</b> KASP Assay by Design für eine Titin-Variante.....	45
<b>Tabelle 9:</b> Genexpressions-Assays.....	45
<b>Tabelle 10:</b> Information der generierten Exom-Sequenzen für HI-Großfamilie..... 2001.....	48
<b>Tabelle 11:</b> Anzahl detektierter Varianten der Exome von HI-Großfamilie 2001 .....	48
<b>Tabelle 12:</b> Geteilte Varianten der Exome von HI-Großfamilie 2001 .....	49
<b>Tabelle 13:</b> Ko-Segregation der identifizierten Varianten in der Familie 2001.....	51
<b>Tabelle 14:</b> Ergebnis des Screenings von <i>GUCY1A3</i> in gesunden Individuen .....	53
<b>Tabelle 15:</b> Ergebnis des Screenings von <i>GUCY1A3</i> in HI-Patienten .....	54
<b>Tabelle 16:</b> Information der generierten Exom-Sequenzen für DCM-Familie 9800..	58
<b>Tabelle 17:</b> Anzahl ausgewählter Varianten der DCM-Familie 9800.....	58
<b>Tabelle 18:</b> Paraloge Genregionen der DCM-Familie 9800 .....	60
<b>Tabelle 19:</b> Geteilte Varianten der Exome von DCM-Familie 9800.....	60
<b>Tabelle 20:</b> Geteilte Varianten der Exome von DCM-Familie 9800 (modifiziert) .....	61
<b>Tabelle 21:</b> Information der generierten Exom-Sequenzen für DCM-Familie..... 16725.....	63
<b>Tabelle 22:</b> Anzahl ausgewählter Varianten der DCM-Familie 16725.....	64
<b>Tabelle 23:</b> Paraloge Genregionen der DCM-Familie 16725 .....	65
<b>Tabelle 24:</b> Geteilte Varianten der Exome von DCM-Familie 16725, die nicht in..... ID 29134 vorkamen .....	66
<b>Tabelle 25:</b> Ko-Segregation der <i>TTN</i> -Variante in der DCM-Familie 16725 .....	68
<b>Tabelle 26:</b> Information der generierten Exom-Sequenzen für DCM-Familie..... 009-1229 .....	70
<b>Tabelle 27:</b> Anzahl ausgewählter Varianten der DCM-Familie 009-1229.....	71
<b>Tabelle 28:</b> Paraloge Genregionen der DCM-Familie 009-1229 .....	72

---

<b>Tabelle 29:</b> Geteilte Varianten der Exome von DCM-Familie 009-1229.....	
(modifiziert/nicht modifiziert) .....	73
<b>Tabelle 30:</b> Ko-Segregation von sechs Varianten in der DCM-Familie 009-1229 ....	77
<b>Tabelle 31:</b> Zusätzlich geteilte Varianten der Exome von DCM-Familie 009-1229.....	
(nicht modifiziert) .....	80
<b>Tabelle 32:</b> Ko-Segregation von drei Varianten in der DCM-Familie 009-1229.....	81
<b>Tabelle 33:</b> Zusätzlich geteilte Varianten der Exome von DCM-Familie 009-1229.....	
(modifiziert).....	82
<b>Tabelle 34:</b> Information der generierten Exom-Sequenzen für HCM-Familie 757....	85
<b>Tabelle 35:</b> Anzahl detektierter Varianten der HCM-Familie 757 .....	86
<b>Tabelle 36:</b> Paraloge Genregionen der HCM-Familie 757 .....	87
<b>Tabelle 37:</b> Geteilte Varianten der Exome von HCM-Familie 757 .....	88
<b>Tabelle 38:</b> Ko-Segregation der <i>HOOK1</i> -Variante in der HCM-Familie 757.....	89
<b>Tabelle 39:</b> Varianten, die in mehr Familien detektiert wurden .....	91
<b>Tabelle 40:</b> Charakterisierung der Probanden der DCM-Familie 16725 .....	127
<b>Tabelle 41:</b> Charakterisierung der Probanden der DCM-Familie 009-1229 .....	128
<b>Tabelle 42:</b> Charakterisierung der Probanden der DCM-Familie 9800 .....	129
<b>Tabelle 43:</b> Charakterisierung der Probanden der HCM-Familie 757 .....	130
<b>Tabelle 44:</b> Charakterisierung der Probanden der HI-Großfamilie 2001 .....	131
<b>Tabelle 45:</b> Charakterisierung des Kollektivs für das Screening von <i>GUCY1A3</i> ....	132
<b>Tabelle 46:</b> Charakterisierung des Kollektivs für die Sequenzierung dreier Gene .	132
<b>Tabelle 47:</b> Oligonukleotide für Ko-Segregation in Familie 2001 .....	133
<b>Tabelle 48:</b> Oligonukleotide für IDAHO-Screening <i>GUCY1A3</i> .....	133
<b>Tabelle 49:</b> Oligonukleotide für Sequenzierung von <i>ETFDH</i> , <i>CCT7</i> und <i>GCLC</i> .....	134
<b>Tabelle 50:</b> Oligonukleotide für Ko-Segregation in Familie 009-1229 und 757 .....	135
<b>Tabelle 51:</b> Parameter von SAMtools für Filterung der Varianten .....	135
<b>Tabelle 52:</b> Anzahl detektierter Varianten in DCM-Familie 9800.....	137
<b>Tabelle 53:</b> Anzahl detektierter Varianten in DCM-Familie 16725.....	139
<b>Tabelle 54:</b> Anzahl detektierter Varianten in DCM-Familie 009-1229 .....	141
<b>Tabelle 55:</b> Anzahl detektierter Varianten in HCM-Familie 757.....	142

## Spezielle Abkürzungen der Arbeit

bp: Basenpaar

BWA: Burrows-Wheeler-Aligner

DCM: dilatative Kardiomyopathie

Gb: Gigabasen =  $10^9$  Basen

GWAS: genomweite Assoziationsstudie

HCM: hypertrophische Kardiomyopathie

HDL: *high-density* Lipoprotein

HI: Herzinfarkt

HRM: *high resolution melting*

IGV: Integrative Genomics Viewer

Indel: Insertion/Deletion

Kb: Kilobasen =  $10^3$  Basen

KHK: koronare Herzerkrankung

LDL: *low-density* Lipoprotein

MAF: *minor allele frequency*

Mb: Megabasen =  $10^6$  Basen

SBS: *Sequencing By Synthesis*

SNP: Einzelnukleotidpolymorphismus

SNV: Einzelnukleotidvariante

UTR: *untranslated region*

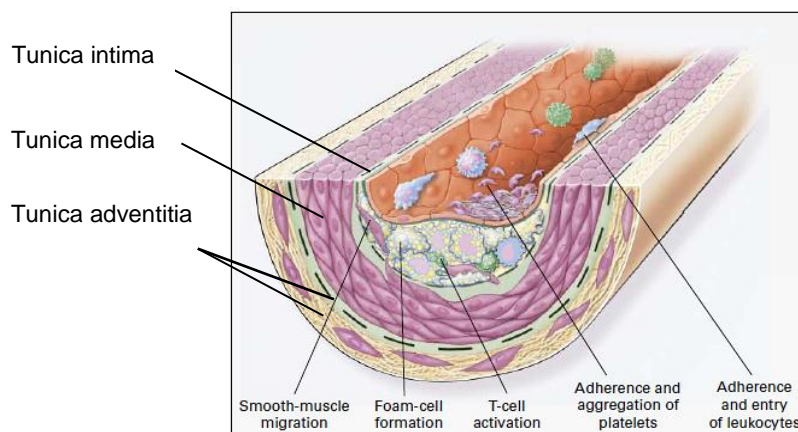
# 1. Einleitung

## 1.1. Epidemiologie kardiovaskulärer Erkrankungen

Kardiovaskuläre Erkrankungen stellen die häufigste Todesursache in der westlichen Welt dar (Roger *et al.*, 2012). Im Jahr 2008 sind in Amerika 82,6 Millionen Menschen kardiovaskulär erkrankt. Diese Erkrankungen mit einer Prävalenz von 36% führten zu 297,7 Milliarden Dollar an Krankenhausaufwendungen. Als Prävalenz wird dabei die relative Häufigkeit von Krankheitsfällen in einer Bevölkerung während einer bestimmten Zeitdauer bezeichnet. Die koronare Herzerkrankung (KHK), die durch Atherosklerose verursacht wird, ist für den größten Teil der kardiovaskulären Mortalität ursächlich (Wiesner *et al.*, 1999). In Deutschland starben laut des Statistischen Bundesamts im Jahr 2006 fast 67.000 Menschen an Herzinfarkt (HI), einer Manifestation der KHK (Erdmann & Schunkert 2007). Mit einer Prävalenz von 3,1% erlitten 2008 7,9 Millionen Menschen in Amerika einen Herzinfarkt. Die Herzinsuffizienz, die vor allem durch KHK und Herzmuskelerkrankungen, den sogenannten Kardiomyopathien, ausgelöst wird, führt mit einer Prävalenz von 2,4% zu einer hohen kardiovaskulären Mortalität (Roger *et al.*, 2012). In Amerika litten 2008 5,7 Millionen Menschen an einer Herzinsuffizienz. Herztransplantationen, die durch Kardiomyopathien notwendig werden, führen in Amerika zu zusätzlich 200 Millionen Dollar an Krankenhausaufwendungen pro Jahr (Towbin & Bowles 2002). Dilatative Kardiomyopathie (DCM) und hypertrophische Kardiomyopathie (HCM) stellen die größten Gruppen der Kardiomyopathien dar. DCM kommt bei Erwachsenen häufiger bei Männern als bei Frauen vor. Die Prävalenz für DCM in der Allgemeinbevölkerung liegt bei 1 zu 2500 (Jefferies & Towbin 2010), die für HCM liegt bei 1 zu 500, wobei Männer doppelt so oft betroffen sind wie Frauen (Harvey & Leinwand 2011).

## 1.2. Atherosklerose und Herzinfarkt

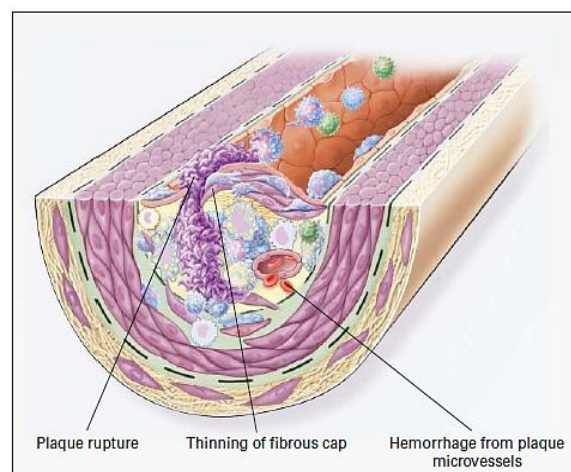
Atherosklerose wird als eine entzündliche Erkrankung der Gefäße angesehen, die durch Interaktion von modifizierten Lipoproteinen, Makrophagen, T-Zellen und weiteren zellulären Elementen der Arterienwand entsteht. Durch eine Fehlfunktion des Endothels, die durch Bluthochdruck und Diabetes begünstigt wird, kommt es zu einer Hochregulation von Adhäsionsmolekülen (z.B. VCAM-1) in der Gefäßwand und dadurch zur Anheftung und Einwanderung von Monozyten und T-Zellen (Ross 1999). Hauptursache von Atherosklerose ist eine hohe Plasmakonzentration von Cholesterin, hauptsächlich von *low-density* Lipoprotein (LDL) Cholesterin. LDL-Partikel dienen zum Transport von Cholesterin zu peripheren Geweben, werden dort über den LDL-Rezeptor auf den Zellen erkannt und aufgenommen. In den Arterien können LDL-Partikel zum Beispiel durch Oxidation (oxLDL) modifiziert werden. oxLDL kann nicht mehr von dem LDL-Rezeptor erkannt werden, sondern wird von dem sog. *scavenger*-Rezeptor, der auf der Oberfläche von aus Monozyten differenzierte Makrophagen und glatten Muskelzellen exprimiert wird, gebunden. Die Aufnahme von oxLDL führt zur Bildung von Lipidperoxiden und zur Anhäufung von Cholesterin in den Zellen (Lipidtröpfchen) wodurch Schaumzellen entstehen (Ross 1999; Glass & Witztum 2001). Es kommt zur Bildung eines sog. *fatty streak* (Fetteinlagerung in der Gefäßwand). *Fatty streaks* sind Anzeichen einer Atherosklerose in der frühen Phase (Abbildung 1).



**Abbildung 1: Bildung von *fatty streaks* in Atherosklerose. Ein *fatty streak* entsteht im Endothelium der Gefäßwand und besteht aus glatten Muskelzellen, Schaumzellen und T-Zellen. Außerdem kommt es zur Anlagerung von Blut-Plättchen und Leukozyten ans Endothel. Abbildung modifiziert nach (Ross 1999).**

Ein *fatty streak* besteht ursprünglich aus Schaumzellen und T-Zellen. Später kommt es zum Einwandern von glatten Muskelzellen. In einem *fatty streak* werden T-Zellen z.B. durch Interleukin-2 und Tumornekrosefaktor  $\alpha$  aktiviert. Weiterhin kommt es zur Migration, Anlagerung und Aggregation von Blut-Plättchen und Leukozyten an die endotheliale Gefäßwand (Abbildung 1). Diese wird durch mehrere Faktoren, wie Stickstoffmonoxid, Selektine und oxLDL stimuliert (Ross 1999; Glass & Witztum 2001).

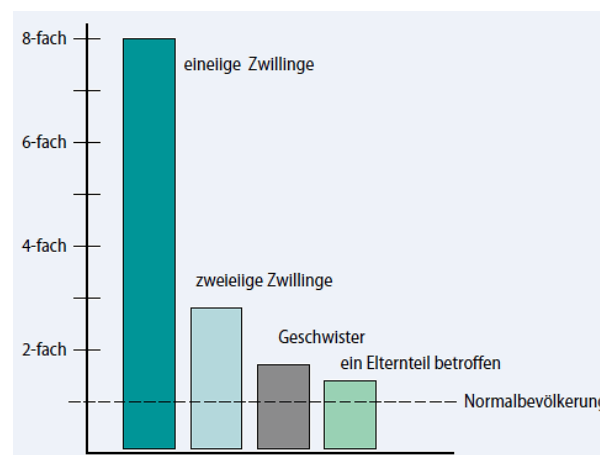
Im fortschreitenden Verlauf der Atherosklerose kommt es in einem *fatty streak* zur Proliferation und weiteren Migration der glatten Muskelzellen, Makrophagen und T-Zellen und damit zur Ausbreitung der Läsion. Diese verändert sich auch strukturell, indem eine fibröse Kappe gebildet wird, die die Läsion vom Lumen abschirmt. Es handelt sich in diesem Stadium um eine fortgeschrittene und komplizierte Läsion (Plaque). Abbildung 2 zeigt die Ruptur eines Plaques, die zu einem Gefäßverschluss führen kann. Die Ruptur tritt gewöhnlich an der Stelle des Plaques auf, wo durch die von den Makrophagen freigesetzten proteolytischen Enzyme ein Abbau der fibrösen Kappe herbeigeführt wurde. Die Enzyme können auch zum Abbau der Matrix und damit zu einer Blutung führen, was wiederum nach Aktivierung der Gerinnungskaskade in einem Gefäßverschluss resultiert (Ross 1999; Glass & Witztum 2001).



**Abbildung 2: Ruptur eines Plaques. Durch von Makrophagen freigesetzte proteolytische Enzyme kommt es zum Abbau der fibrösen Kappe und der Matrix. Dadurch kann Ruptur und Blutung auftreten (Ross 1999).**

Zu einem Herzinfarkt kommt es bei Verschluss einer Herzkranzarterie. Durch die Durchblutungsstörung wird der betroffene Bereich des Herzens (Infarktgebiet) nicht mehr mit Sauerstoff versorgt und stirbt ab (Erdmann & Schunkert 2007).

Risikofaktoren für die Entstehung von Atherosklerose können unter anderem Hypertonie, Diabetes mellitus, Hyperlipidämie und Rauchen sein (Ross 1999). Zudem wurde festgestellt, dass eine positive Familienanamnese ein Risikofaktor für die Entstehung von kardiovaskulären Erkrankungen ist (Abbildung 3). In einer Studie mit eineiigen und zweieiigen Zwillingen wurde die Risikoerhöhung eines gesunden Zwillingspartners untersucht, wenn der andere an Herzinfarkt verstorben war. Es zeigte sich, dass eineiige Zwillinge eine Risikoerhöhung um den Faktor 8,1 und zweieiige Zwillinge eine Risikoerhöhung um den Faktor 2,6 haben (Marenberg *et al.*, 1994; Erdmann & Schunkert 2007). In anderen Studien wurde gezeigt, dass eine Risikoerhöhung von 1,5 – 1,8 besteht, wenn ein Elternteil betroffen ist. Etwas höher liegt dieser Wert, wenn ein Bruder oder eine Schwester betroffen ist (Andresdottir *et al.*, 2002; Horne *et al.*, 2006; Erdmann & Schunkert 2007).



**Abbildung 3: Risikoerhöhung für KHK/MI in Anbetracht der familiären Disposition. Bei eineiigen Zwillingen: Risikoerhöhung 8-fach, bei zweieiigen Zwillingen: Risikoerhöhung 2,6-fach. Bei Erkrankung von Eltern oder Geschwistern: Risikoerhöhung 1,5 - 1,8-fach (Erdmann & Schunkert 2007).**

Da die positive Familienanamnese als Risikofaktor für KHK und HI gilt, wurde bislang versucht, die genetischen Komponenten zu identifizieren. Mehrere Gene sowie Umweltfaktoren spielen vermutlich bei der Entstehung von KHK und HI eine Rolle. Es handelt sich hierbei um komplexe Erkrankungen, die keinem Mendel-Erbgang folgen. Dabei sollte nicht vergessen werden, dass Risikofaktoren wie Hypertonie und



Diabetes mellitus ebenfalls durch eine genetische Komponente beeinflusst werden (Erdmann & Schunkert 2007). Besonders deutlich wird das bei monogener Hyperlipidämie, die durch Defekte im LDL-Rezeptor Stoffwechselweg ausgelöst wird und sehr häufig zur Ausprägung einer KHK/HI führt. Kausale Mutationen für Hyperlipidämie wurden in den Genen *LDLR*, *APOB*, *PCSK9* und *LDLRAP1* identifiziert (Rader *et al.*, 2003; Soutar 2010). In einem ersten Ansatz zur Identifizierung der genetischen Komponente von KHK/HI wurden Kandidatengen-Studien durchgeführt. Diese beruhten auf der Hypothese, dass Proteine, die in der Pathogenese von Atherosklerose beteiligt sind, Varianten in den zugrunde liegenden Genen tragen, die die Pathogenese von KHK ebenso beeinflussen (Schunkert *et al.*, 2010). Häufige Polymorphismen in Kandidatengenen wurden auf Assoziation mit KHK/HI untersucht. Dieser Ansatz wurde als Fall-Kontroll-Studie durchgeführt, die Individuen genotypisiert und die Allelfrequenzen von nicht verwandten Fällen und Kontrollen verglichen (Abbate *et al.*, 2008). Erstmals im Jahr 1992 konnte ein Zusammenhang zwischen einem Polymorphismus im Gen *angiotensin-converting enzyme (ACE)* mit Herzinfarkt nachgewiesen werden (Cambien *et al.*, 1992). In den folgenden Jahren wurden viele Studien durchgeführt und z.B. Assoziationen von KHK/HI mit Polymorphismen in den Genen *APOE* und *PAI* und mehreren Koagulationsfaktoren gezeigt (Wilson *et al.*, 1994; Doggen *et al.*, 1998; Gardemann *et al.*, 1999). Trotz einer Vielzahl an Veröffentlichungen konnte die Assoziation der meisten Kandidatengene nicht repliziert werden. Unter anderem zu kleine Studienpopulationen und Heterogenität in den Studien führten zu falsch positiven Assoziationsbefunden bei KHK/HI (Schunkert *et al.*, 2010).

Neben den Kandidatengen-Studien wurde mit einem anderen Ansatz versucht die Genetik von KHK und HI zu entschlüsseln (Erdmann & Schunkert 2007). In seltenen Fällen folgt das Auftreten von KHK/HI einem scheinbar autosomal-dominanten Vererbungsmuster (Mendel-Erbgang). Familien mit gehäuft auftretenden Krankheitsfällen, die ein solches Vererbungsmuster vermuten ließen, wurden in mehreren Studien untersucht. Mit Mikrosatelliten, die über das Genom verteilt liegen, wurden alle Familienmitglieder genotypisiert. Es wurde vermutet, dass bestimmte Allele häufiger mit der Krankheit auftreten (ko-segregieren), als zu erwarten wäre (Abbate *et al.*, 2008). Die sog. Kopplungsanalyse (*linkage analysis*) führte zur Identifizierung von mehreren chromosomalen Regionen, die mit KHK und HI ein positives Kopplungssignal zeigten (Broeckel *et al.*, 2002; Samani *et al.*, 2005). Im

Jahr 2003 wurde durch eine Kopplungsanalyse in einer Familie auf Chromosom 15q26 ein positives Signal identifiziert. In der Region befanden sich ungefähr 93 Gene, wobei nur das Gen *MEF2A* als plausibles Kandidatengen für KHK/HI angesehen wurde. Nach Sequenzierung von *MEF2A* in den Familienmitgliedern wurde eine Deletion von 21 Basenpaaren in allen betroffenen Familienmitgliedern identifiziert (Wang *et al.*, 2003). In der Arbeit von Anna Helgadottir *et al.* wurde im Jahr 2004 durch ein Kopplungssignal ein bestimmter Haplotyp im Gen *ALOX5AP* einem höheren Risiko für Herzinfarkt zugeordnet. Der Haplotyp umfasste vier Einzelnukleotidpolymorphismen (*single-nucleotide polymorphism*, SNP) in dem Gen (Helgadottir *et al.*, 2004). Kopplungsanalysen in Familien mit KHK/HI erwiesen sich aber generell als schwierig. Einerseits können aufgrund der Häufigkeit der Krankheit in der Allgemeinbevölkerung Phänokopien in den Familien enthalten sein. Man spricht von einer Phänokopie, wenn die Krankheit durch andere Faktoren und nicht durch einen familiären Gendefekt verursacht wurde. Bei einer angenommenen Lebenszeitprävalenz von etwa 2,5% (Wiesner *et al.*, 1999) sind in großen Stammbäumen einige Phänokopien anzunehmen. Außerdem erwies sich die genaue Zuordnung der Gene in den Regionen und die Sequenzierung dieser Gene zur Identifizierung von Mutationen als mühevoll. Bislang wurden in den Kopplungsanalysen private Mutationen identifiziert, die nur in der untersuchten Familie die Entstehung von KHK/HI erklärten (Schunkert *et al.*, 2010). Zum Beispiel konnte die in der oben beschriebenen Arbeit von Lejin Wang *et al.* gezeigte Kausalität von *MEF2A* durch Sequenzierung des Gens in nicht verwandten KHK-Patienten oder in HI-Familien nicht mehr gezeigt werden (Weng *et al.*, 2005; Lieb *et al.*, 2008). Genotypisierung beschriebener SNPs von *MEF2A* in nicht verwandten HI-Patienten sowie das Ergebnis einer Fall-Kontroll-Assoziationsstudie, in der die Assoziation eines häufigen Polymorphismus in *MEF2A* mit KHK untersucht werden sollte, konnten die Kausalität nicht untermauern (Lieb *et al.*, 2008; Liu *et al.*, 2012).

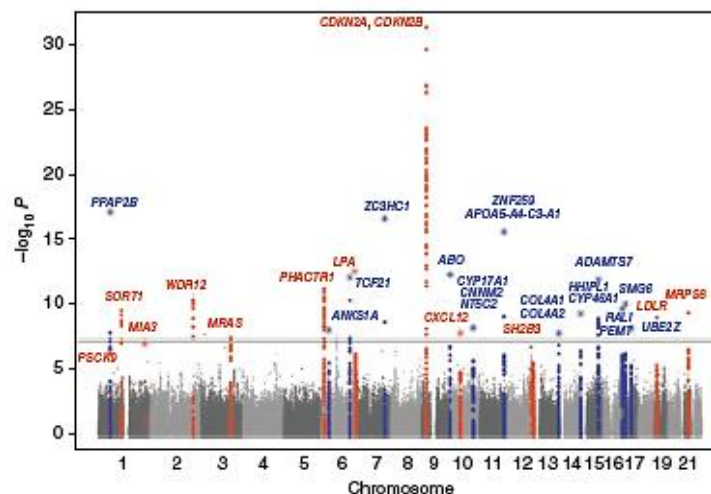
Ein Durchbruch für die Studie von komplexen Erkrankungen wurde durch die Entschlüsselung der Sequenz des menschlichen Genoms und der gefolgten Annotation der SNPs im Genom ermöglicht. Als SNPs werden Einzelnukleotid-Variationen des Genoms bezeichnet, die sehr häufig sind (Schunkert *et al.*, 2010; O'Donnell & Nabel 2011). Mit Arrays, die Tausende bis Millionen an SNPs detektieren, konnten Individuen im hohen Durchsatz genotypisiert werden (Schunkert *et al.*, 2010). Es wurde die Hypothese aufgestellt, dass eine häufig auftretende

Erkrankung durch SNPs mit einer häufigen Allelfrequenz von mindestens 5% in der Allgemeinbevölkerung assoziiert ist. Mit dieser sog. *common disease – common variant* Hypothese wurden genomweite Assoziationsstudien (*genomewide association study*, GWAS) durchgeführt (Schunkert *et al.*, 2010; O'Donnell & Nabel 2011; Shields 2011). Bei diesem Ansatz wurden in einem Fall-Kontroll-Design Individuen genotypisiert (Schunkert *et al.*, 2010). Im Jahr 2007 wurden die Ergebnisse der ersten GWAS zu KHK und HI veröffentlicht, wobei ein Locus auf Chromosom 9p21 in allen Studien das stärkste Assoziationssignal zeigte (McPherson *et al.*, 2007; Helgadottir *et al.*, 2007; WTCCC 2007). Weitere Loci, die eine Assoziation mit KHK/HI zeigten, wurden identifiziert (Samani *et al.*, 2007; Kathiresan *et al.*, 2009; Tregouet *et al.*, 2009; Erdmann *et al.*, 2009). In Tabelle 1 sind GWAS für häufige kardiovaskuläre Erkrankungen aufgelistet (O'Donnell & Nabel 2011).

**Tabelle 1: GWAS von häufigen kardiovaskulären Erkrankungen (O'Donnell & Nabel 2011)**

Study	Clinical Outcome	Sample Size		Major Ethnic Group	Selected Major Findings
		Genomewide Association	Replication		
<i>number of subjects</i>					
Kathiresan et al., 2009 <sup>13</sup>	Myocardial infarction	6,042	19,492	European	9 Loci, including 9p21 and <i>SORT1</i>
Schunkert et al., 2011 <sup>14</sup>	Coronary artery disease	86,995	56,682	European	>20 Loci, including <i>ABO</i> and <i>ADAMTS7</i>
C4D Consortium, 2011 <sup>15</sup>	Coronary artery disease	30,482	40,593	European, East Asian	Novel loci, including <i>LIPA</i>
Smith et al., 2010 <sup>16</sup>	Heart failure incidence	23,821	NA	European, African	Suggestive loci, including <i>LRIG3</i>
Morrison et al., 2010 <sup>17</sup>	Heart failure mortality	2,992	NA	European, African	Suggestive loci, including <i>ADAMTS12</i>
Villard et al., 2011 <sup>18</sup>	Dilated cardiomyopathy	2,287	2,467	European	Top loci: <i>HSPB7</i> and <i>BAG3</i>
Ikram et al., 2009 <sup>19</sup>	Ischemic stroke	19,602	7,269	European, African	Top locus: <i>NINJ2</i>
Yamada et al., 2009 <sup>20</sup>	Ischemic stroke	267	5,981	Japanese	Top locus: <i>CELSR1</i>
Akiyama et al., 2010 <sup>21</sup>	Intracranial aneurysm	482	1,398	Japanese	Suggestive loci, including <i>ARHGGEF11</i> and <i>TMEM195</i>
Bilguvar et al., 2008 <sup>22</sup>	Intracranial aneurysm	7,856	1,171	European, Japanese	Top loci: 9p21 and <i>SOX17</i>
Yasuno et al., 2010 <sup>23</sup>	Intracranial aneurysm	15,295	4,777	European, Japanese	Associations with 9p21 and <i>CNNM2</i>
Koriyama et al., 2010 <sup>24</sup>	Peripheral arterial disease	1,553	2,239	Japanese	Top loci: <i>OSBPL10</i> and 10p12.31
Thorgeirsson et al., 2008 <sup>25</sup>	Peripheral arterial disease	10,995	4,848	European	Top locus: <i>CHRNA3</i>
Elmore et al., 2009 <sup>26</sup>	Abdominal aortic aneurysm	235	1,421		Suggestive locus: 3p12.3
Gretarsdottir et al., 2010 <sup>27</sup>	Abdominal aortic aneurysm	31,795	10,718	European	Top loci: 9p21 and <i>DAB2IP</i>
Benjamin et al., 2009 <sup>28</sup>	Lone atrial fibrillation	40,518	6,218	European	Top loci: <i>PITX2</i> and <i>ZFHX3</i>
Gudbjartsson et al., 2007 <sup>29</sup>	Lone atrial fibrillation	36,137	5,806	European	Top loci: 4q25 and <i>ZFHX3</i>
Ellinor et al., 2010 <sup>30</sup>	Lone atrial fibrillation	14,179	4,771	European	Top loci: <i>KCNN3</i> and 20q13.13
Holm et al., 2011 <sup>31</sup>	Sick sinus syndrome	38,384	1,654	European	Top locus: <i>MYH6</i>
Bezzina et al., 2010 <sup>32</sup>	Ventricular fibrillation	972	537	European	Top locus: <i>CXADR</i>
Arking et al., 2010 <sup>33</sup>	Sudden cardiac arrest	650	19,611	European	Top locus: <i>GPC5</i>
Arking et al., 2011 <sup>34</sup>	Sudden cardiac death	22,055	14,265	European	Top locus: <i>BAZ2B</i>
Tregouët et al., 2009 <sup>35</sup>	Venous thromboembolism	1,647	3,237	European	Top locus: <i>ABO</i>

Beispielsweise wurden bei der im Jahr 2009 veröffentlichten GWA-Studie von Sekar Kathiresan *et al.* 2.967 Herzinfarkt-Patienten und 3.075 gesunde Kontrollen genotypisiert. Die Ergebnisse wurden in einem Kollektiv von 19.492 Individuen repliziert (Tabelle 1). Insgesamt zeigten neun Loci eine signifikante Assoziation (Kathiresan *et al.*, 2009). Um weitere Loci identifizieren zu können, musste die Fallzahl dieser Studien erhöht werden. Deshalb wurde in der Arbeit von Heribert Schunkert *et al.* eine Metaanalyse aus 14 GWAS mit insgesamt 22.233 KHK-Fällen und 64.762 gesunden Kontrollen durchgeführt. Die Replikation erfolgte in einem Kollektiv aus 56.682 Individuen. In dieser Arbeit wurden 13 neue Loci identifiziert, die mit KHK ein genomweites Assoziationssignal zeigten. Zehn von zwölf Loci aus früheren Arbeiten konnten ebenso repliziert werden (Tabelle 1). In Abbildung 4 ist der Manhattan-Plot für diese Studie dargestellt. Alle SNPs wurden entsprechend der chromosomalen Lage und des p-Wertes für Assoziation aufgetragen. Ein p-Wert von 0,05 gilt als signifikant. Da aber das multiple Testen der unabhängigen SNPs einbezogen werden muss, wurde für die genomweite Signifikanz eine Schwelle bei einem p-Wert von  $5 \cdot 10^{-8}$  festgelegt (Dudbridge & Gusnanto 2008). SNPs mit einem geringeren p-Wert, d.h. die über der Signifikanzschwelle liegen, zeigten eine signifikante Assoziation mit KHK (Schunkert *et al.*, 2011).



**Abbildung 4: Manhattan-Plot der Ergebnisse der CARDIoGRAM-Metaanalyse. Auf der Abszisse sind die Chromosomen aufgetragen, auf der Ordinate  $-\log_{10}P$  für alle SNPs. Die graue Linie markiert genomweite Signifikanz. Gene mit einem signifikant assoziierten SNP sind dargestellt in rot für bereits bekannte Loci und in blau für neu entdeckte Loci (Schunkert *et al.*, 2011).**

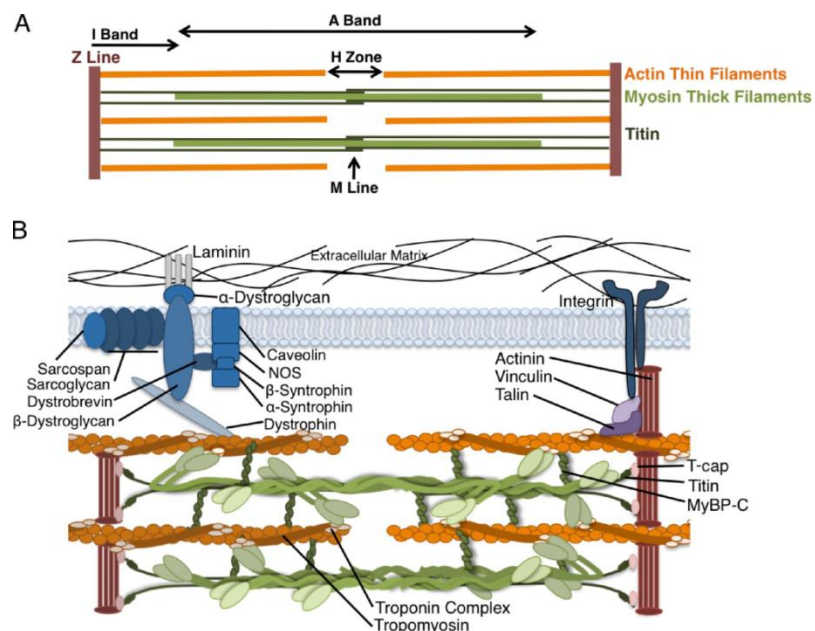
---

Insgesamt wurden in GWAS mehr als 30 Loci identifiziert, die mit HI und KHK eine Assoziation zeigen. Für HI zeigte der Locus auf Chromosom 9p21 das stärkste Assoziationssignal. Der assoziierte SNP liegt in einer Region ohne bekannte Proteinkodierende Gene. Vielmehr wird das nicht-kodierende RNA Molekül *ANRIL* der Region zugeordnet. Loci, die mit KHK assoziiert sind, betreffen, wie erwartet, auch Gene, die für Änderungen im Lipidstoffwechsel sorgen, wie *SORT1*, *PCSK9*, *HNF1A* und *MRAS*. Trotz der Beschreibung einer Vielzahl an Loci bleibt ihre Entdeckung häufig ohne klinische Relevanz. Jede Variante leistet nur einen kleinen Beitrag bei der Entstehung der Erkrankungen (O'Donnell & Nabel 2011). Außerdem sind die assoziierten SNPs der Studien oft nicht kausal für die Erkrankung. Die Identifizierung der kausalen Varianten, die dem Assoziationssignal zugrunde liegen, bedarf weiterer Forschung (O'Donnell & Nabel 2011).

Mit den GWAS wurde die *common disease – common variant* Hypothese verfolgt (Shields 2011). Die durch GWAS identifizierten Varianten erklären aber nur 13% der Heritabilität von KHK (Peden & Farrall, 2011), die auf insgesamt 60% geschätzt wird (Schnabel et al., 2012). Die *odds ratios* dieser Varianten lagen bei 1,02 – 1,92 und zeigten somit nur schwache Effekte (Peden & Farrall 2011). Als *odds ratio* wird die Wahrscheinlichkeit einer Risikoerhöhung angegeben, wenn man Träger des Risikoallels ist. Die Frage stellte sich nach der fehlenden Heritabilität dieser komplexen Erkrankung. Bei der monogenen Hyperlipidämie wurden z.B. seltene Varianten mit starkem Effekt in den Genen *PCSK9* und *LDLR* beschrieben (Manolio et al., 2009). Seltene Varianten mit einer Allelfrequenz <0,5% wurden durch GWAS noch nicht untersucht und könnten die fehlende Heritabilität von KHK erklären (Manolio et al., 2009; Peden & Farrall 2011).

### 1.3. Kardiomyopathien

Kardiomyopathien sind schwerwiegende Erkrankungen des Herzmuskels, die durch zelluläre Fehlfunktionen entstehen und zu Herzinsuffizienz führen können (Harvey & Leinwand 2011). Das Herz ist auf ein Netzwerk von Zellen angewiesen, um seine Funktion aufrechtzuerhalten. Das Netzwerk der Zellen besteht aus Kardiomyozyten, die kontrahierenden Muskelzellen im Herzen, Endothelzellen, glatten Muskelzellen, Fibroblasten und Immunzellen. Die Kardiomyozyten sind elektrophysiologisch durch *gap junctions* verbunden. Da die Weiterleitung von Impulsen nur über die *gap junctions* erfolgt, ist die Anzahl und Verteilung dieser wichtig für die Geschwindigkeit der Impulsweiterleitung (Saffitz 2005). *Gap junctions* koordinieren die Kontraktion einzelner Kardiomyozyten und über die Verbindung zur extrazellulären Matrix die Kontraktion des gesamten Herzens. Die Funktionseinheit von Kardiomyozyten ist das Sarkomer, das aus sich wiederholenden Einheiten von Aktin und Myosin besteht (Abbildung 5A). Das Sarkomer an sich besteht aus mehr als 20 Proteinen und ist über eine Vielzahl weiterer Proteine mit der extrazellulären Matrix gekoppelt (Abbildung 5B). Zelluläre Fehlfunktionen treten auf, wenn das komplexe Netzwerk an Interaktionen gestört wird (Harvey & Leinwand 2011).

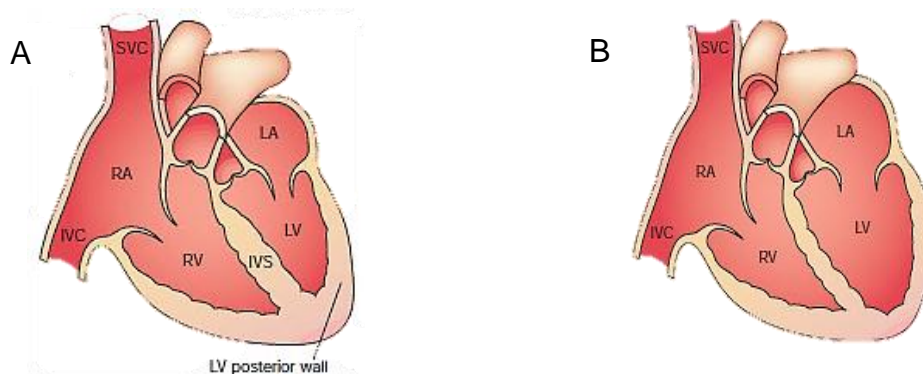


**Abbildung 5: Schematische Darstellung eines Sarkomers. A: Grundorganisation eines Sarkomers. Aktinfilamente sind an der Z-Scheibe verankert und interagieren mit den Myosinfilamenten. Titin verbindet die Z-Scheibe mit der M-Scheibe. Die M-Scheibe, I-Bande und A-Bande werden durch ihre Komponenten definiert. B: Darstellung der wichtigsten Proteine in einem Sarkomer (Harvey & Leinwand 2011).**

Es gibt vier Gruppen von Kardiomyopathien: dilatative Kardiomyopathie (DCM), hypertrophische Kardiomyopathie (HCM), restriktive Kardiomyopathie und arrhythmogen-rechtsventrikuläre Kardiomyopathie (Towbin & Bowles 2002). DCM und HCM stellen die größten Gruppen dar und werden im Weiteren genauer erläutert.

### 1.3.1. Dilatative Kardiomyopathie

Die DCM ist eine Herzmuskelerkrankung, die durch Erweiterung des linken Ventrikels und Verdünnung der Ventrikelwand beschrieben wird (Abbildung 6B). Dadurch kommt es zu einem fortschreitenden Verlust der Auswurfleistung (Ejektionsfraktion, EF) und zur eingeschränkten Kontraktilität (Towbin & Bowles 2002).



**Abbildung 6: Darstellung von einem gesunden Herzen/Herzen mit DCM. A: Gesundes Herz. B: Herz mit DCM charakterisiert durch erweiterten linken Ventrikel und dünner Ventrikelwand. IVS: interventrikuläres Septum; LA: linkes Atrium; LV: linker Ventrikel; RA: rechtes Atrium; RV: rechter Ventrikel; SVC: Vena cava superior; IVC: Vena cava inferior (Towbin & Bowles 2002).**

Neben der Dilatation des linken Ventrikels und der systolischen Fehlfunktion können auch Rückfluss des Blutes durch die Mitralklappe (*mitral regurgitation*) und ventrikuläre Arrhythmien auftreten. Ungefähr 30% – 40% aller DCM-Fälle sind familiär bedingt. Bei dieser sogenannten familiären DCM liegt oft ein autosomal-dominantes Vererbungsmuster vor. X-chromosomal und autosomal-rezessive Vererbung tritt nur in seltenen Fällen auf (Towbin & Bowles 2002).



Die erste Mutation wurde 1993 durch Kopplungsanalysen in männlichen Geschwistern mit X-chromosomal vererbter DCM im Dystrophin-Gen (*DMD*) identifiziert (Towbin *et al.*, 1993). In dieser Arbeit wurde gezeigt, dass die Krankheit durch eine starke Reduktion oder Abwesenheit des Dystrophin-Proteins im Herzen verursacht wurde (Towbin *et al.*, 1993; Towbin & Bowles 2002). In Familien mit autosomal-dominanter DCM wurden über Kopplungsanalysen mehrfach Mutationen im *TTN*-Gen (*giant-muscle filament titin*) identifiziert. Titin ist ein wichtiges Strukturprotein im Sarkomer (vgl. Abbildung 5) und spielt dadurch eine entscheidende Rolle bei der Kontraktion (Gerull *et al.*, 2002; Gerull *et al.*, 2006).

Neben den Kopplungsanalysen wurden auch Sequenzierungen von Kandidatengen in Patienten mit idiopathischer DCM (Ursache der Erkrankung nicht bekannt) oder familiärer DCM durchgeführt. Im Jahr 1998 wurden dadurch erstmals Mutationen im kardialen Aktin-Gen (*CACT*, *ACTC*) als kausal für autosomal-dominant vererbte DCM beschrieben. Durch diese Mutationen wurde eine veränderte Kraftübertragung während der Kontraktion vermutet (Olson *et al.*, 1998). Weitere kausale Mutationen wurden unter anderem in den Genen *TNNT2* (kardiales Troponin T), *SCN5A* (kardialer Natriumkanal), *MYH7* (schwere Kette von  $\beta$ -Myosin) und in *ANKRD1* (*Ankyrin repeat domain 1*) identifiziert (Hershberger *et al.*, 2008; Duboscq-Bidot *et al.*, 2009). In Tabelle 2 sind die Gene, die als kausal für DCM beschrieben wurden, aufgrund ihrer Ontologie in verschiedenen Gruppen zusammengefasst (Hershberger & Siegfried 2011). Kausale Gene für DCM gehören hauptsächlich in Gruppen, die für Sarkomer- oder Zytoskelett-Proteine kodieren. Zusätzlich wurden mehrere Gene identifiziert, die für Z-Scheiben Proteine kodieren (Jefferies & Towbin 2010).

**Tabelle 2: Ontologie kausaler DCM-Gene (Hershberger & Siegfried 2011)**

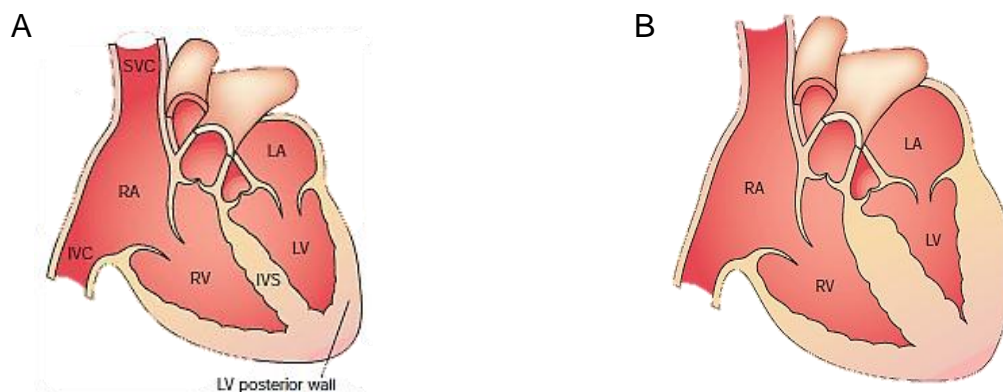
Sarcomere	Z-Disc	Cytoskeleton	Mitochondrial	RNA Binding	Ion Channel	Gamma Secretase Activity	Sarcoplasmic Reticulum	Transcription Factor	Nuclear Envelope
<i>ACTC</i>	<i>TCAP</i>	<i>DMD</i>	<i>TAZ/G4.5</i>	<i>RBM20</i>	<i>ABCC</i>	<i>PSEN1</i>	<i>PLN</i>	<i>EYA4</i>	<i>LMNA</i>
<i>MYH7</i>	<i>CSR3P</i>	<i>DES</i>			<i>SCN5A</i>	<i>PSEN2</i>			<i>TMPO</i>
<i>MYH6</i>	<i>ACTN2</i>	<i>LDB3</i>							
<i>MYBPC3</i>	<i>MYPN</i>	<i>SGCD</i>							
<i>TNNT2</i>	<i>ANKRD1</i>	<i>PDLIM3</i>							
<i>TNNC1</i>		<i>VCL</i>							
<i>TNNI3</i>		<i>RYAB</i>							
<i>TPM1</i>		<i>ILK</i>							
<i>TTN</i>		<i>LAMA4</i>							



Die Häufigkeit von bekannten DCM-Mutationen wurde als sehr gering eingeschätzt. In nur 30% - 35% aller familiären DCM-Fälle wurde eine genetische Ursache identifiziert. Zudem wurde die Genetik der idiopathischen DCM diskutiert (Hershberger & Siegfried 2011). Um letztere genauer zu untersuchen, wurde eine umfassende Assoziationsstudie mit idiopathischen DCM-Fällen durchgeführt (Stark *et al.*, 2010). In dieser Arbeit wurde gezeigt, dass ein Polymorphismus im Gen *HSBP7*, kodierend für ein kardiales *heat shock*-Protein, mit idiopathischer DCM assoziiert ist. Eine weitere GWAS zeigte die Assoziation von Polymorphismen in den Genen *HSPB7* und *BAG3* mit Herzinsuffizienz, die durch DCM verursacht wurde (Villard *et al.*, 2011). Da beide Gene für Proteine kodieren, die für die Degradation anderer Proteine wichtig sind, wurde ein geschädigter Stoffwechselweg für abzubauen Proteine als neuer Mechanismus vorgeschlagen. Trotz der Erkenntnisse der letzten Jahre wurde die Genetik der DCM nur geringfügig entschlüsselt.

### 1.3.2. Hypertrophische Kardiomyopathie

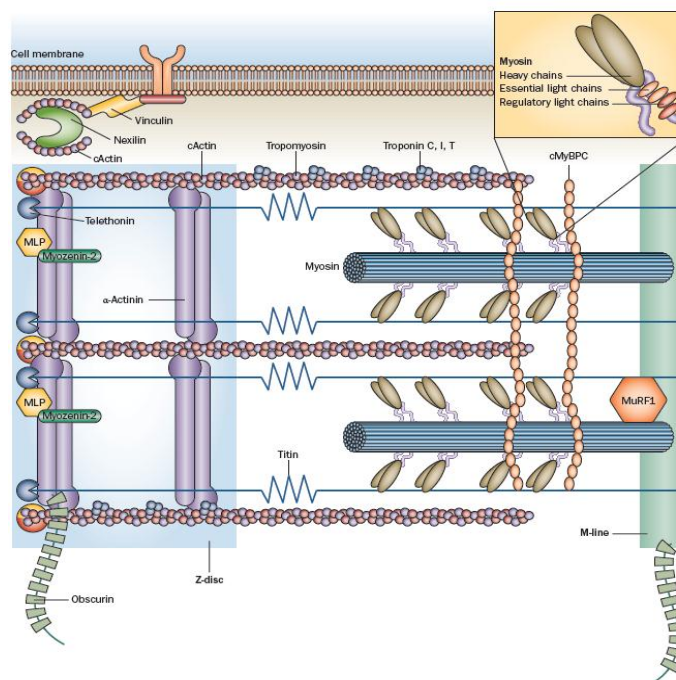
Bei der HCM handelt es sich um eine Herzmuskelerkrankung, die durch eine Verdickung des interventrikulären Septums und der linken Ventrikelwand beschrieben wird (Abbildung 7B). Dadurch kommt es zu einer verbesserten Pumpfunktion (hyperkontraktile systolische Funktion) mit mangelnder Relaxation des Herzens, einer sog. diastolische Fehlfunktion (Towbin & Bowles 2002).



**Abbildung 7: Darstellung von einem gesunden Herzen/Herzen mit HCM. A: Gesundes Herz. B: Herz mit HCM charakterisiert durch verdicktes interventrikuläres Septum und verdickter linker Ventrikelwand. IVS: interventrikuläres Septum; LA: linkes Atrium; LV: linker Ventrikel; RA: rechtes Atrium; RV: rechter Ventrikel; SVC: Vena cava superior; IVC: Vena cava inferior (Towbin & Bowles 2002).**

Eine Behinderung des Blutflusses vom linken Ventrikel zur Aorta kann ebenfalls auftreten (Towbin & Bowles 2002). Durch eine zelluläre Fehlfunktion (vgl. 1.3) kommt es zur Zunahme der Herzmasse und Verdickung der Ventrikelwand. Dies wird, histologisch betrachtet, von einer Vergrößerung der Kardiomyozyten mit auftretender Fehlanordnung der Zell-Zellkontakte und einer Fibrose ausgelöst (Harvey & Leinwand 2011).

Mehr als 50% aller HCM-Fälle sind familiär bedingt. Das Auftreten der Krankheit folgt einem autosomal-dominanten Erbgang und ist dadurch genetisch homogener als DCM (Kimura 2008). Im Jahr 1990 wurde erstmals durch eine Kopplungsanalyse eine kausale Mutation im Gen kodierend für die schwere Kette von kardialem  $\beta$ -Myosin (*MYH7*) identifiziert (Geisterfer-Lowrance *et al.*, 1990). Darauffolgend wurden mehrere kausale Gene und hunderte von Mutationen erkannt. Hauptsächlich kodieren diese Gene für sarkomerische Proteine (Abbildung 8).



**Abbildung 8: Proteine, in denen kausale Mutationen für HCM gefunden wurden. Die meisten Proteine stellen Komponenten des Sarkomers dar (Frey *et al.*, 2012).**

Mutationen in *MYH7* und *MYBP-C* (Myosin-Bindeprotein-C) wurden am häufigsten als kausal für HCM beschrieben und erklären etwa die Hälfte aller HCM-Fälle (Tabelle 3). Ungefähr 10% - 15% aller HCM-Fälle werden durch Mutationen im kardialen Troponin T- (*TNNT2*), kardialen Troponin I- (*TNNI3*) und im  $\alpha$ -

Tropomyosin- (*TPM1*) Gen erklärt (Marian 2010). Seltener Mutationen finden sich im Titin-Gen, das bereits bei DCM als kausal beschrieben wurde (vgl. 1.3.1). Durch Sequenzierung in 82 HCM-Patienten konnte z.B. die Assoziation der Arg740Leu Mutation im Titin-Gen mit HCM gezeigt werden (Sato *et al.*, 1999). In Tabelle 3 sind Gene für HCM aufgelistet, in denen Mutationen unterschiedlich häufig als kausal beschrieben wurden (Marian 2010).

**Tabelle 3: Kausale Gene für HCM. Modifiziert nach (Marian 2010)**

Gen	Symbol	Häufigkeit
<i>β-Myosin heavy chain</i>	<i>MYH7</i>	~25%
<i>Myosin binding protein C</i>	<i>MYBPC3</i>	~25%
<i>Cardiac troponin T</i>	<i>TNNT2</i>	~3% - 5%
<i>Cardiac troponin I</i>	<i>TNNC1, TNNI3</i>	~3% - 5%
<i>α-Tropomyosin</i>	<i>TPM1</i>	~1%
<i>Myozenin 2</i>	<i>MYOZ2</i>	1:250
<i>Myosin light chain 1</i>	<i>MYL3</i>	selten
<i>Myosin light chain 2</i>	<i>MYL2</i>	selten
<i>α-Actin</i>	<i>ACTC1</i>	selten
<i>Titin</i>	<i>TTN</i>	selten
<i>Telethonin</i>	<i>TCAP</i>	selten

Trotz der Aufklärung eines großen Teils der Genetik von HCM, sind noch 25% - 35% aller HCM-Fälle ungeklärt. Die Identifizierung der zugrunde liegenden kausalen Gene mit den Mutationen ist eine zukünftige Herausforderung (Seidman & Seidman 2011).

Durch die Erforschung der Genetik von komplexen Phänotypen wurden in den letzten Jahren neue Erkenntnisse erzielt. Die Hypothese, dass seltene Varianten mit starkem Effekt die fehlende Heritabilität häufiger Erkrankungen erklären können, steht nun im Vordergrund (Marian 2012). Diese sogenannte *common disease – rare variant* Hypothese kann zum jetzigen Zeitpunkt durch GWAS nicht erforscht werden, da die SNP-Arrays nur häufige Allele abdecken (Peden & Farrall 2011). Durch die Entwicklung neuer Sequenzierungsmethoden wird die Erforschung seltener Varianten ermöglicht (Marian 2012).

## 1.4. Geschichte der Sequenzierung

Im Jahr 1977 wurde durch die Arbeit von Frederick Sanger *et al.* die Molekularbiologie revolutioniert (Sanger *et al.*, 1977; Schuster 2008). Es wurde die Sequenzierung des ersten Genoms, das Genom eines Bakteriophagen, durch **Kettenabbruch-Sequenzierung** beschrieben (Sanger *et al.*, 1977). Diese Methode wurde 1986 durch Verwendung von Fluorophor-markierten Dideoxynukleotiden in einem Reaktionsansatz verbessert (Smith *et al.*, 1986). Den vorläufigen Höhepunkt der Sequenzierungs-Ära lieferte das Humangenomprojekt, in dem eine vollständige Sequenz des menschlichen Genoms durch Sanger-Sequenzierung (3000 Mb) erstellt wurde (Lander *et al.*, 2001; Venter *et al.*, 2001). Da die klassische Sanger-Sequenzierung einige Einschränkungen aufweist, unter anderem die geringe Anzahl an Proben, die gleichzeitig analysiert werden können, wurde die Entwicklung neuer Techniken anvisiert. In einem Patent von EMBL wurde 1991 die Sequenzierung durch Synthese (*Sequencing By Synthesis*) vorgestellt, bei der Fluorophor-markierte Basen durch eine sensitive CCD-Kamera während der Synthese detektiert werden (Ansorge & EMBL Heidelberg 1991). Ein Gel oder ein anderes Separationsmedium für die Fragmente ist nicht mehr nötig. Das vorgestellte Prinzip führte zur Entwicklung der **Hochdurchsatz-Sequenzierung** (*next-generation sequencing*), bei der Sequenzierungen massiv parallel ablaufen können. Diese Methode wurde von der Zeitschrift *Nature Methods* als „Methode des Jahres 2007“ deklariert (Ansorge 2009). Die neuen Entwicklungen führten unter anderem zur Bildung eines internationalen Konsortiums, das *1000 Genomes Project*, das die Sequenzierung von 1000 menschlichen Genomen plante. Dadurch sollte ein neuer Katalog als Referenz für menschliche Variation im Genom entstehen (Kaiser 2008; Siva 2008). Durch die Entwicklung der Hochdurchsatz-Sequenzierung sind Kosten und Dauer für Sequenzierungen deutlich gefallen. Die Sequenzierung bietet eine Vielzahl von Möglichkeiten, um die fehlende Heritabilität von Krankheiten aufzuklären. Zum Beispiel diente die **Exom-Sequenzierung** in den letzten zwei Jahren als Ansatz um neue kausale Mutationen in seltenen Krankheiten zu identifizieren (Ku *et al.*, 2012). Krankheiten, die einem klassischen Mendel-Erbgang folgen, werden häufig durch Mutationen in Protein-kodierenden Regionen verursacht. Das Exom umfasst alle Protein-kodierenden Regionen des Genoms. Mit circa 18.000 - 19.000 Genen stellt das Exom ungefähr 1% - 2% (~ 30 Mb - 60 Mb) des menschlichen Genoms dar (Hershberger & Siegfried 2011). Die Sequenzierung des Exoms erhöht somit die

Wahrscheinlichkeit Varianten mit einem starken Effekt zu entdecken. Eine erste Arbeit wurde von Sarah Ng *et al.* vorgestellt, in der zwölf Exome von Menschen sequenziert wurden (Ng *et al.*, 2009). In der Arbeit wurden unter anderem vier nicht verwandte Individuen sequenziert, die an dem Freeman-Sheldon-Syndrom leiden. Das zugrunde liegende Gen war bereits bekannt, so dass der Ansatz als *proof-of-concept* angesehen wurde. Alle Varianten pro Exom wurden identifiziert und anschließend gefiltert. Bereits bekannte Varianten aus der Datenbank dbSNP und aus den damals acht abgeschlossenen HapMap-Exomen wurden ausgeschlossen. Verbleibende Varianten wurden als neue Varianten bezeichnet und zwischen den sequenzierten Individuen verglichen. Es wurden nicht-synonyme Varianten, Spleiß-Varianten und Insertionen und Deletionen bevorzugt, die laut einer bioinformatischen Vorhersage als schädlich für die entstehenden Proteine eingestuft wurden. Diese Filterung der Varianten führte zur Identifizierung einer Mutation im Gen *MYH3*, die von allen vier betroffenen Individuen geteilt wurde (Ng *et al.*, 2009). In einer zweiten Arbeit von Sarah Ng *et al.* wurde von vier nicht verwandten Individuen, die an dem Miller-Syndrom leiden, das Exom sequenziert. Filterung der Varianten und Vergleich zwischen den Individuen führte zur Identifizierung einer Mutation im Gen *DHODH*, die als kausal für Miller-Syndrom betrachtet wurde (Ng *et al.*, 2010b). In weiteren Arbeiten wurden durch Exom-Sequenzierungen Mutationen im Gen *SETBP1* als kausal für Schinzel-Giedion-Syndrom, Mutationen im Gen *WDR35* als kausal für das Sensenbrenner-Syndrom und Mutationen im Gen *MLL2* als kausal für das Kabuki-Syndrom beschrieben (Hoischen *et al.*, 2010; Gilissen *et al.*, 2010; Ng *et al.*, 2010a).

Zur Identifizierung kausaler Varianten wurden drei Kriterien in allen Arbeiten berücksichtigt: Erstens wurden häufige Varianten ausgeschlossen. Grundlage für diese Filterung sind öffentliche Datenbanken, wie dbSNP oder das *1000 Genomes Project* oder von den Arbeitsgruppen selbst sequenzierte Kontrollen. Zweitens wurden Varianten mit starkem Effekt bevorzugt, d.h. nicht-synonyme Varianten, Spleiß-Varianten und Insertionen/Deletionen im kodierenden Bereich. Drittens wurde der funktionelle Effekt der Varianten mit bioinformatischen Vorhersage-Programmen (z.B. PolyPhen) bestimmt und nur die als schädlich vorhergesagten Varianten genauer betrachtet. Je nach Studiendesign wurden weitere Kriterien gesetzt. Zum Beispiel wurden für rezessiv vererbte Krankheiten homozygote Varianten und für dominant vererbte Krankheiten heterozygote Varianten bevorzugt. Wenn betroffene Familienmitglieder sequenziert wurden, wurden die Daten verglichen, um die kausale

Variante zu identifizieren, die von den Familienmitgliedern geteilt wird. Oftmals gingen dabei auch Ergebnisse aus Kopplungsanalysen ein, um den Bereich für die Identifizierung von Varianten einzuschränken (Ku *et al.*, 2012).

Es gibt mehrere Hochdurchsatz-Sequenzierungsplattformen, wobei für diese Arbeit die Plattform von Illumina genutzt wurde. Die genaue Methode ist unter 3.5.1 und 3.5.2 beschrieben.

## 1.5. Zielsetzung

In dieser Arbeit sollten die genetischen Ursachen für Herzinfarkt, DCM und HCM untersucht werden. Wie bereits beschrieben, wurden für KHK/HI bereits mögliche Kandidatengene identifiziert, die allerdings nur geringfügig die Heritabilität dieser komplexen Erkrankung erklären. Für DCM und HCM wurden kausale Mutationen beschrieben, wobei HCM häufiger durch monogene Mutationen verursacht wird als DCM. Die Genetik dieser komplexen Erkrankungen ist aber noch nicht vollständig entschlüsselt. Es wurde die Hypothese aufgestellt, dass seltene Varianten (Frequenzen in der Allgemeinbevölkerung  $<0,5\%$ ) mit starkem Effekt die fehlende Heritabilität von HI, DCM und HCM aufklären sollen. Um die seltenen Varianten zu identifizieren, wurden Familien mit gehäuft auftretenden Krankheitsfällen ausgewählt. Das Auftreten der Krankheit folgte in diesen Familien einem klassischen Mendel-Erbgang mit autosomal-dominantem Vererbungsmuster. Die seltenen Varianten sollten von den betroffenen Familienmitgliedern gegenüber den nicht betroffenen Familienmitgliedern geteilt werden, so dass die Variante vollständig mit der Krankheit ko-segregiert. Zur Identifizierung dieser kausalen Varianten sollte je nach Stammbaum das Exom von zwei bis fünf betroffenen Familienmitgliedern sequenziert werden. Für die Exom-Sequenzierungen sollten für die Erstellung der Exom-Bibliotheken zwei verschiedene Ansätze (Agilent Technologies und Illumina) verwendet werden. Außerdem sollte die Auswertung und Filterung der Varianten mit unterschiedlichen Strategien und Software erfolgen. Wie in den bereits veröffentlichten Arbeiten zur Exom-Sequenzierung dargestellt, sollten bekannte Varianten ausgeschlossen werden. Von den verbleibenden Varianten sollten nicht-synonyme Varianten, Spleiß-Varianten und Insertionen/Deletionen im kodierenden Bereich bevorzugt werden. Da ein dominantes Vererbungsmuster in den Familien angenommen wurde, sollten zudem heterozygote Varianten ausgewählt werden. Nach Filterung und Identifizierung der möglich kausalen Varianten, sollten diese mit PolyPhen-2 und Konservierungsanalysen als schädlich oder nicht schädlich für die entstehenden Proteine eingestuft werden. Außerdem sollte die Kausalität der Varianten durch Ko-Segregation in den weiteren betroffenen und nicht betroffenen Familienmitgliedern überprüft werden. Dadurch sollten neue kausale Mutationen für Herzinfarkt, DCM und HCM entdeckt werden, die die Pathophysiologie dieser Krankheiten weiter aufklären.

## 2. Material

### 2.1. Geräte

ABI PRISM® 7000 Sequence Detection System	Applied Biosystems, Foster City, USA
ABI PRISM® 7900HT Sequence Detection System	Applied Biosystems, Foster City, USA
Agilent 2100 Bioanalyzer	Agilent Technologies, Santa Clara, USA
Bioruptor mit 0.5 & 0.65 ml tube holder	Diagenode, Liège, Belgien
cBot™ v 1.5	Illumina, San Diego, USA
Covaris S-series Single Tube Sample Preparation System, Model S2	Covaris, Woburn, USA
3730 DNA Analyzer	Applied Biosystems, Foster City, USA
Freezer -20°C	Liebherr, Bulle, Schweiz
Genome Analyzer IIx	Illumina, San Diego, USA
HiSeq2000	Illumina, San Diego, USA
LightScanner®	Idaho Technology Inc., Salt Lake City, USA
Mastercycler pro S	Eppendorf, Hamburg, Deutschland
NanoDrop 1000 Spectrophotometer	Peqlab, Erlangen, Deutschland
TeMoRoboter	Tecan, Männedorf, Schweiz



## 2.2. Verbrauchsmaterialien

Agilent DNA 1000 Kit	Agilent Technologies, Santa Clara, USA
Agilent High Sensivity DNA Kit	Agilent Technologies, Santa Clara, USA
AmpliTaq® Gold DNA Polymerase	Applied Biosystems, Foster City, USA
Covaris micro TUBE with AFA fiber and snap cap	Covaris, Woburn, USA
6x DNA Loading Dye	Fermentas, Vilnius, Lithuania
GeneRuler™ 100bp DNA Ladder	Fermentas, Vilnius, Lithuania
Human Cardiac Myocytes, cryopreserved HCM-c	PromoCell, Heidelberg, Deutschland
Human MTC Panel I	Clontech, Mountain View, USA
KASPar 5000 reagentkit	KBioscience, Hoddesdon, England
KASP by Design	KBioscience, Hoddesdon, England
LCGreen® Plus und Melting Dye	Idaho Technology Inc., Salt Lake City, USA
0.65 ml microtubes for DNA shearing	Diagenode, Liège, Belgien
MinElute Gel Extraction Kit	Qiagen, Hilden, Deutschland
MinElute PCR Purification Kit	Qiagen, Hilden, Deutschland
Myocyte Growth Medium	PromoCell, Heidelberg, Deutschland
Puregene DNA purification system	Qiagen, Hilden, Deutschland
SureSelect Human All Exon Kit v.1.0.1	Agilent Technologies, Santa Clara, USA

---

TaqMan® Gene Expression Assay	Applied Biosystems, Foster City, USA
TaqMan® Gene Expression Mastermix	Applied Biosystems, Foster City, USA
TruSeq™ DNA Sample Preparation Kit	Illumina, San Diego, USA
TruSeq™ Exome Enrichment Kit	Illumina, San Diego, USA
Paired-End Genomic DNA Sample Prep Kit	Illumina, San Diego, USA
TruSeq™ Paired end Cluster Kit v.2.5 HS	Illumina, San Diego, USA
TruSeq™ SBS-Kit HS	Illumina, San Diego, USA

### 2.3. Software

CASAVA v.1.8	Illumina, San Diego, USA
Cyrillic	Cyrillic Software, Wallingford, England
Integrative Genomics Viewer	Broad Institute, Cambridge, USA
LightScanner™ Software v.1.1.0.566	Idaho Technology Inc., Salt Lake City, USA
Reference Manager 12	Thomson Reuters, New York City, USA
SDS v.2.3	Applied Biosystems, Foster City, USA
SeqScape v.2.7	Applied Biosystems, Foster City, USA
Sequence Scanner v.1.0	Applied Biosystems, Foster City, USA

## 3. Methoden

### 3.1. Rekrutierung

#### 3.1.1. Rekrutierungsstrategie Herzinfarkt-Großfamilien

Im Rahmen der Deutschen Herzinfarkt-Familienstudie wurden am Universitätsklinikum Regensburg Großfamilien mit gehäuft auftretendem Herzinfarkt und koronarer Herzerkrankung rekrutiert. Patienten, die vor ihrem sechzigsten Lebensjahr einen Herzinfarkt erlitten haben (Index-Patienten) und zumindest einen Verwandten haben, der an Herzinfarkt oder schwerer koronarer Herzerkrankung leidet, wurden in die Studie aufgenommen. Die genaue Rekrutierungsstrategie ist anderweitig beschrieben (Fischer *et al.*, 2005). Konnten mindestens fünf betroffene Familienmitglieder rekrutiert werden, handelte es sich um Großfamilien. So konnten 26 Großfamilien mit insgesamt 1001 Personen in sogenannte Studie eingeschlossen werden.

#### 3.1.2. Rekrutierungsstrategie EUROGENE-Heart Failure Project

Das Universitätsklinikum Regensburg nahm zudem an dem EUROGENE-Heart Failure Project, das aus mehreren europäischen Zentren gebildet wurde, teil. Ziel der Studie war die Erkennung von Risikofaktoren und Ursachen von Herzmuskel-erkrankungen. Es wurden Familien mit dilatativer Kardiomyopathie und mit hypertrophischer Kardiomyopathie rekrutiert. Bislang wurden insgesamt 100 Familien mit DCM und 250 Familien mit HCM in dem Projekt eingeschlossen. Die genaue Rekrutierungsstrategie und bisherige Ergebnisse des Projektes sind anderweitig beschrieben (Friedrich *et al.*, 2009; Duboscq-Bidot *et al.*, 2009).

### **3.1.3. Auswahl der Familien für diese Arbeit**

Für diese Arbeit wurden Familien mit gehäuft auftretenden Krankheitsfällen ausgewählt, in denen das Auftreten der Krankheit einem autosomal-dominanten Vererbungsmuster folgt. Es sollten mindestens fünf Individuen, am besten Verwandte zweiten Grades betroffen sein, um optimale Voraussetzungen für Ko-Segregationsanalysen zu gewährleisten. Eine Ausnahme stellt die DCM-Familie 009-1229 dar, die in dieser Arbeit aufgrund des Vorliegens eines schweren Krankheitsbildes (vgl. Tabelle 41) untersucht wurde. Die Stammbäume wurden mit der Software Cyrillic (Cyrillic Software, Wallingford, England) erstellt.

## **3.2. Molekularbiologische Methoden**

### **3.2.1. DNA-Isolierung und Konzentrationsbestimmung**

Genomische DNA aus Gesamtblut wurde mit dem Puregene DNA purification system (Qiagen, Hilden, Deutschland) nach Angaben des Herstellers extrahiert und bei -20°C gelagert. Die DNA-Konzentration wurde mit dem NanoDrop 1000 Spectrophotometer (Peqlab, Erlangen, Deutschland) bestimmt.

### **3.2.2. Oligonukleotide**

Oligonukleotide wurden bei der Firma Eurofins MWG Operon (Huntsville, USA) bestellt und sind im Anhang unter 8.3 Tabelle 47 – 50 mit weiteren Angaben zu den PCR-Bedingungen aufgelistet.

### 3.2.3. Amplifikation von DNA-Fragmenten mittels PCR

Die Polymerasekettenreaktion (PCR) ermöglicht die enzymatische Vervielfältigung von DNA-Fragmenten *in-vitro* (Saiki *et al.*, 1985; Mullis & Faloona 1987). Dies geschieht durch zyklisches Wiederholen des Schmelzens der doppelsträngigen DNA (Denaturierung), gefolgt von Anlagerung (*annealing*) und Verlängerung (Elongation) zweier Primer (synthetische Oligonukleotide), die auf gegenüberliegenden Seiten und in entgegengesetzter Richtung das zu amplifizierende DNA-Fragment flankieren. Dabei erfolgt die Amplifikation des gewünschten DNA-Stücks exponentiell. Bei Amplifikation aus genomischer DNA wurde die AmpliTaq® Gold DNA Polymerase (Applied Biosystems, Foster City, USA) verwendet. Die Reaktion wurde standardmäßig in 20 µl Volumen in einem Thermocycler (Eppendorf, Hamburg, Deutschland) bei 105°C Deckeltemperatur durchgeführt. Die Primermenge betrug je 500 pmol. Die Reaktion wurde mit 20 ng - 100 ng zu amplifizierender DNA-Matrize, 1 U thermostabiler Polymerase und 250 nM dNTPs in Reaktionspuffer standardmäßig wie folgt durchgeführt:

1. Denaturierung 95°C	9 min
2. Denaturierung 95°C	15 sec
3. <i>Annealing</i> TA°C	15 sec
4. Elongation 72°C	x sec
5. finale Elongation 72°C	7 min
6. Abkühlung 10°C	∞

Die Schritte zwei bis vier wurden zyklisch 35- bis 40-mal wiederholt. Die *annealing* Temperatur (TA) ergab sich aus der Schmelztemperatur der Oligonukleotide und wird mit einer Gradienten-PCR (PCR mit verschiedenen *annealing* Temperaturen) ermittelt. Die Elongationszeit ergab sich aus der Länge des zu amplifizierenden Produktes (je 30 sec Elongation für 500 bp PCR-Produkt).

### 3.2.4. Touchdown-PCR

Bei der Touchdown-PCR wird die Amplifikation mit einem modifizierten Programm durchgeführt. Dadurch wird verhindert, dass unspezifische Produkte amplifiziert werden. Folgendes Programm wurde verwendet:

1. Denaturierung 95°C	9 min	
2. Denaturierung 95°C	15 sec	
3. <i>Annealing</i> 72°C	15 sec	-1°C pro Zyklus
4. Elongation 72°C	x sec	

Die Schritte zwei bis vier wurden zyklisch wiederholt bis die entsprechende *annealing* Temperatur erreicht wurde.

5. Denaturierung 95°C	15 sec
6. <i>Annealing</i> TA°C	15 sec
7. Elongation 72°C	x sec

Für 33 Zyklen wurden anschließend die Schritte fünf bis sieben wiederholt.

8. finale Elongation 72°C	7 min
9. Abkühlung 10°C	∞

### 3.2.5. Agarosegelelektrophorese

Die Agarosegelelektrophorese wurde zur Analyse von DNA-Fragmenten nach einer PCR verwendet. Zur Herstellung der Gele wurde 1% Agarose in TBE-Puffer (10x TBE-Puffer: 270 g Tris, 137,5 g Borsäure, 18,6 g EDTA auf 2,5 l H<sub>2</sub>O) durch Aufkochen gelöst. Nach Abkühlung auf etwa 60°C wurden pro ml Gellösung 0,2 µl der 10 mg/ml Stammlösung des Fluoreszenzfarbstoffes Ethidiumbromid zugesetzt und die Agaroselösung in die Gelkammer gegossen. Beim Abkühlen entsteht ein Gel mit netzartiger Struktur, in dem durch Anlegen einer Spannung DNA der Größe nach aufgetrennt werden kann. Die DNA-Proben wurden mit 6x DNA Loading Dye (Fermentas, Vilnius, Lithuania) verdünnt und in die Geltaschen pipettiert. Als Größenstandard wurden 5 µl des GeneRuler™ 100 bp DNA Ladder (Fermentas) verwendet. Die Elektrophorese erfolgte 40 min bei 100 V und 85 mA. Unter UV-Licht wurden die DNA-Fragmente mit eingelagertem Ethidiumbromid sichtbar (Sharp *et al.*, 1973). Die Dokumentation der Gele erfolgte über einen UV-Transilluminator mit angeschlossener CCD-Kamera (Eigenherstellung).

### **3.2.6. Isolierung und Aufreinigung von DNA-Fragmenten**

Zur Isolierung aus einem Agarosegel wurden die Banden mit einem sterilen Skalpell unter UV-Licht ausgeschnitten und mit dem MinElute Gel Extraction Kit (Qiagen) nach Angaben des Herstellers aufgereinigt. Die Aufreinigung von PCR-Produkten ohne Isolierung aus einem Agarosegel erfolgte mit dem MinElute PCR Purification Kit (Qiagen).

### **3.2.7. DNA-Sequenzierung**

Die Verifizierung von Nukleotidsequenzen von amplifizierten PCR-Produkten erfolgte über Auftragssequenzierung bei der Firma GENEART (Regensburg, Deutschland). Ein Sequenzieransatz bestand standardmäßig aus 15 ng - 20 ng DNA je 100 bp PCR-Fragment und 1 µl (10 µM) Primer. Des Weiteren wurde in Kooperation am Institut für Humangenetik an der Universität Bonn Sequenzierungen im Plattenformat durchgeführt. Die Auswertung der Sequenzen erfolgte mit Sequence Scanner v.1.0 (Applied Biosystems, Foster City, USA).

### **3.2.8. Gelelektrophorese mit Bioanalyser**

Der Agilent 2100 Bioanalyser (Agilent Technologies, Santa Clara, USA) sorgt für eine präzise Analyse von DNA mit hoher Sensitivität. Die Proben werden elektrophoretisch getrennt und durch einen eingelagerten Fluoreszenzfarbstoff detektiert. Im Vergleich zur Agarosegelelektrophorese ist zudem der Probenverbrauch minimiert. DNA-Fragmentgrößen und DNA-Konzentrationen wurden mit dem Agilent High Sensitivity DNA Kit und Agilent DNA 1000 Kit (Agilent Technologies) bestimmt. Entsprechend der Sensitivitätsgrenze der Kits wurden unterschiedliche DNA-Mengen eingesetzt ([www.agilent.com](http://www.agilent.com)).

### 3.3. Exom-Sequenzierung

Für die Exom-Sequenzierungen wurden für die Erstellung der Exom-Bibliotheken die Verbrauchsmaterialien von zwei verschiedenen Herstellern, Agilent Technologies und Illumina, verwendet. Die Sequenzierungen wurden mit der Hochdurchsatz-Sequenzierungsplattform von Illumina, mit den Geräten Genome Analyzer IIx und HiSeq2000, durchgeführt. Bei der anschließenden Auswertung wurde, getrennt nach Anreicherung der Exom-Bibliotheken und Sequenziermaschine, mit unterschiedlicher Software gearbeitet.

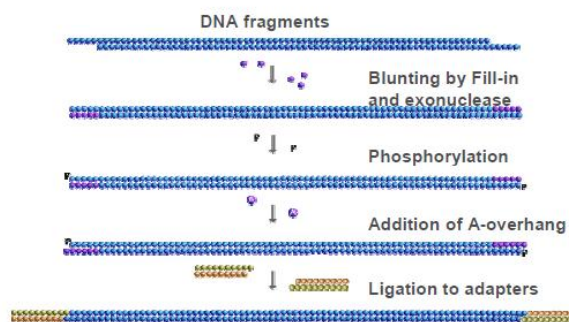
Im Bereich des *next-generation sequencing* wurden viele englische Ausdrücke ins Deutsche übernommen. Zur Erstellung der Exom-Bibliothek werden die Exone mit Hilfe von komplementären Sonden abgedeckt. Bei diesem sog. *exon capturing* können damit exonische Sequenzen vom restlichen Genom isoliert und anschließend sequenziert werden. Während der Sequenzierung werden *sequence reads* der Exone generiert. Dabei handelt es sich um Sequenzen, die von einem einzigen DNA-*Template* erzeugt werden. Bei dem *mate-pair sequencing* (oder *paired-end sequencing* ([www.illumina.com](http://www.illumina.com))) wird jedes DNA-*Template* dabei sowohl vom 5`- als auch vom 3`-Ende aus sequenziert. Anschließend müssen die erhaltenen *sequence reads* mit einer Referenzsequenz verglichen und zugeordnet werden. Dieser Vorgang wird als *mapping* bezeichnet. Wie oft jede einzelne Base sequenziert und der Referenzsequenz zugeordnet wird, wird dabei mit der *coverage* ausgedrückt (Lupski *et al.*, 2010).



### 3.4. Erstellung der Exom-Bibliotheken

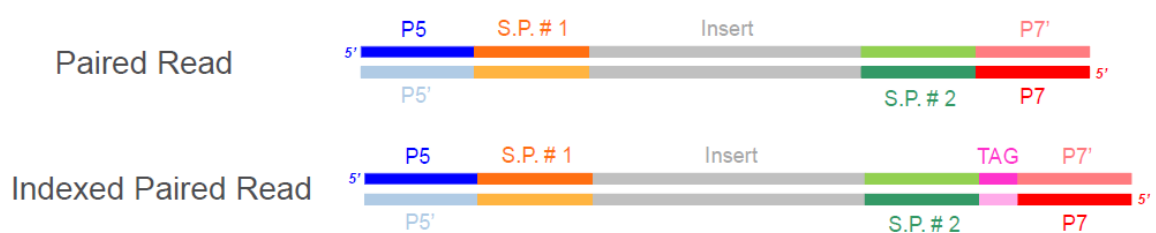
#### 3.4.1. Erstellung der Exom-Bibliothek nach Agilent Technologies

Für die Anreicherung der Exom-Anteile aus den Genomen wurde der SureSelect Human All Exon Kit (Agilent Technologies) mit Reagenzien von New England Biolabs (Ipswich, USA) verwendet. Im ersten Schritt wurden zur Erstellung der Bibliotheken 3 µg der genomischen DNA mit dem Covaris System (Covaris, Woburn, USA) fragmentiert. Anders als im Protokoll beschrieben, wurde die DNA für fünf Zyklen je 30 sec mit Ultraschall behandelt. Die erhaltenen Enden der 200 bp DNA-Fragmente wurden mittels der T4-DNA-Ligase/Polymerase und dem Klenow-Enzym repariert und die 5`-Enden mit der T4-Polynukleotidkinase phosphoryliert. Bevor die Ligation spezifischer Adapter an die DNA-Fragmente erfolgen konnte, wurde am 3`-Ende ein Desoxyadenosinphosphat angehängt (Abbildung 9).



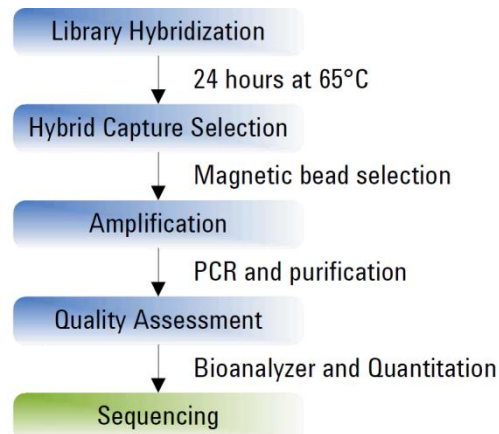
**Abbildung 9: Erstellung einer Bibliothek für die Exom-Sequenzierung – Teil 1.** Die DNA-Fragmente werden repariert und phosphoryliert, bevor ein Desoxyadenosinphosphat am 3`-Ende angehängt werden kann. Danach erfolgt die Ligation der Adapter ([www.illumina.com](http://www.illumina.com)).

Eine Adaptersequenz bestand aus der Bindestelle für die *flowcell* und der Bindestelle für ein Oligonukleotid (Abbildung 10 *Paired Read*).



**Abbildung 10: Adapter für *paired-end* Sequenzierung mit und ohne Index.** P5 und P7: Bindestelle für *flowcell*, S.P.#1 und S.P.#2: Bindestelle für Primer, TAG: Indexsequenz ([www.illumina.com](http://www.illumina.com)).

Überschüssige, nicht ligierte Adapter wurden abgetrennt und die Bibliotheken mittels PCR amplifiziert. Die Bibliotheken wurden anschließend mit der *Sure Select Biotinylated RNA library* vermischt. Bei dieser Hybridisierungsreaktion binden exonische Sequenzen an die biotinylierte Bibliothek. Über Streptavidin-Kügelchen konnten diese Fragmente dann gebunden und nochmals amplifiziert werden. Die Aufreinigung der Produkte nach Adapterligation und Amplifikation der Exom-Bibliothek erfolgte mittels *solid phase reverse immobilization technology* (SPRI-Kügelchen). Dabei wurde für die Elution der DNA EB-Puffer (10 mM Tris-Cl, pH 8,5) und nicht H<sub>2</sub>O verwendet (Abbildung 11).



**Abbildung 11: Erstellung einer Bibliothek für die Exom-Sequenzierung – Teil 2. Hybridisierung mit der *Sure Select Biotinylated RNA library*. Es erfolgt eine Aufreinigung der Exom-Bibliothek mit Streptavidin-Kügelchen ([www.agilent.com](http://www.agilent.com)).**

### 3.4.2. Erstellung der Exom-Bibliothek nach Illumina und Vergleich

Für die Exom-Sequenzierungen mit dem TruSeq™ DNA Sample Preparation Kit/TruSeq™ DNA Exome Enrichment Kit von Illumina (San Diego, USA) wurden alle benötigten Reagenzien von dem Hersteller mitgeliefert. Das Prinzip der Anreicherung der Exom-Anteile aus dem Genom entsprach dem Anreicherungsprinzip von Agilent Technologies (vgl. 3.4.1). Allerdings wurde im 96 *well*-Plattenformat gearbeitet und nicht in Einzelreaktionsgefäßen. Weitere Unterschiede der Anwendungen sind in Tabelle 4 zusammengefasst.

**Tabelle 4: Vergleich der Erstellung der Exom-Bibliotheken**

	<b>Agilent Technologies</b>	<b>Illumina</b>
<i>Target</i>	38 Mb	62 Mb
Ausgangsmaterial	3 µg genomische DNA	1 µg genomische DNA
Fragmentierung	Covaris	Bioruptor
Optimale Fragmentgröße	150 bp - 200 bp	200 bp - 400 bp
Aufreinigung nach Fragmentierung	PCR purification kit	-
Qualitätsbestimmung	Bioanalyzer	Bioanalyzer (nicht im Protokoll)
Aufreinigung nach Reparatur der Enden	PCR purification kit	AMPure <i>beads</i> (SPRI-Kügelchen)
Aufreinigung nach Anhängen von Desoxyadenosinphosphat	PCR purification kit	-
Ligation der Adapter	-	+ Ligation der Indices
Aufreinigung nach Ligation	AMPure <i>beads</i> (SPRI-Kügelchen)	AMPure <i>beads</i> (SPRI-Kügelchen) Agarosegelelektrophorese
Qualitätsbestimmung	Bioanalyzer	-
Amplifikation	6 Zyklen	10 Zyklen
Aufreinigung nach Amplifikation	PCR purification kit	AMPure <i>beads</i> (SPRI-Kügelchen)
Validierung der Bibliothek	Bioanalyzer	Bioanalyzer
<i>Pooling</i> der Bibliotheken und Hybridisierung	Kein Vermischen der DNA 500 ng DNA, Hybridisierung: 24 h	DNA 3-fach/6-fach vermischt Je 500 ng pro Bibliothek, 1. Hybridisierung: 16 - 20 h 2. Hybridisierung: 16 - 20 h
Aufreinigung nach Hybridisierung	Streptavidin-Kügelchen, PCR purification kit	Streptavidin-Kügelchen
Amplifikation nach Hybridisierung	10 - 12 Zyklen	10 Zyklen
Aufreinigung nach Amplifikation	AMPure <i>beads</i> (SPRI-Kügelchen)	AMPure <i>beads</i> (SPRI-Kügelchen)
Qualitätsbestimmung	Bioanalyzer	Bioanalyzer
Cluster Erzeugung	cBot, DNA-Konzentration pro Bibliothek 8 pM	cBot, DNA-Konzentration pro Bibliothek 8 pM

Die Fragmentierung der DNA wurde in speziellen Reaktionsgefäßen für Beschallung (Diagenode, Liège, Belgien) mit dem Bioruptor (Diagenode) durchgeführt. Dafür wurde die DNA mit EB-Puffer (10 mM Tris-Cl, pH 8,5) auf eine Konzentration von 20 ng/µl verdünnt und in einem Volumen von 100 µl fragmentiert. Vor Fragmentierung wurde die DNA 15 min auf Eis gekühlt. Anschließend erfolgte in siebenmaliger Wiederholung eine Beschallung von 30 sec bei niedriger Energie (*low power*) gefolgt von 30 sec ohne Beschallung. Um die DNA wieder im Reaktionsgefäß zu sammeln, wurde kurz zentrifugiert, bevor nochmal acht Zyklen gestartet wurden. Nach den insgesamt 15 Zyklen der Beschallung hatte ein Großteil der DNA eine durchschnittliche Fragmentgröße von 200 bp – 400 bp.

Anders als beim SureSelect Human All Exon Kit ist der TruSeq™ Exome Enrichment Kit dafür geeignet, verschiedene DNA-Proben vor der Hybridisierung mit der Exom-Bibliothek zu vermischen (*pooling* der DNA-Proben). Eine Indexsequenz in der Adaptersequenz führt dazu, dass bei der Ligation der Adapter jeder Probe ein bestimmter Index zugewiesen wird (Abbildung 10 *Indexed Paired Read*). Illumina bietet bis zu zwölf verschiedene Indices an, so dass maximal zwölf Proben vermischt werden konnten. Nach der Ligation wurden die Produkte auf ein Agarosegel geladen. Dadurch konnte DNA mit der optimalen Fragmentgröße (300 bp - 500 bp; optimale Größe nach Fragmentierung 200 bp - 400 bp + Länge der Adapter und Indices 100 bp) aus dem Gel ausgeschnitten und aufgereinigt werden. Ein weiterer wichtiger Unterschied fand sich bei den Hybridisierungszeiten. In dem SureSelect Human All Exon Kit war eine Hybridisierung von 24 h vorgesehen, bei dem TruSeq™ Exome Enrichment Kit handelte es sich um zwei Hybridisierungsschritte (Tabelle 4).

### **3.4.3. Modifiziertes Protokoll zur Erstellung der Exom-Bibliothek nach Illumina**

Für die Exom-Bibliotheken wurden die DNA-Proben in zwei Ansätzen für die Sequenzierung vorbereitet. Ein Ansatz wurde entsprechend dem Protokoll von Illumina durchgeführt. Beim zweiten Ansatz wurde das Protokoll leicht verändert, indem der Schritt der Größenselektion der Fragmente durch ein Agarosegel ausgelassen wurde. Für die anschließende Amplifikation der DNA wurde die Zyklenzahl auf vier reduziert.

## 3.5. Hochdurchsatz-Sequenzierungsplattform Illumina

### 3.5.1. Cluster Erzeugung für Sequenzierung

Die vorbereiteten Exom-Bibliotheken werden auf eine sogenannte *flowcell* gegeben (Abbildung 12). Eine *flowcell* besteht aus acht Kapillaren, auf deren Oberfläche sich die komplementären Sequenzen der Adapter befinden (*Flowcell-Adaptersequenz*). Für die Erzeugung von Cluster (Cluster = Mehrzahl an identischen DNA-Strängen) wird die *flowcell* in eine Maschine, den cBot, geladen (Abbildung 12).

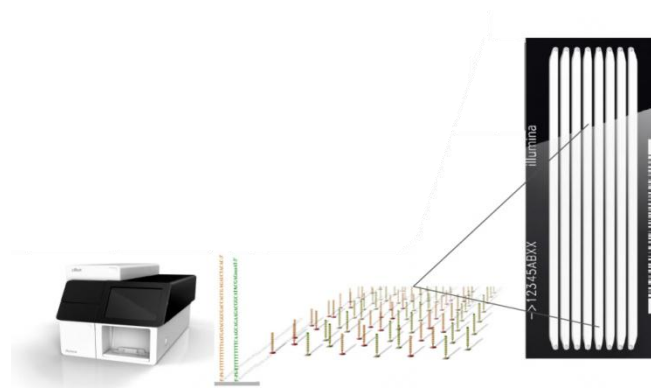


Abbildung 12: cBot und *flowcell*. Für die Erzeugung der Cluster wird der cBot verwendet. Auf der *flowcell*, bestehend aus acht Kapillaren, befinden sich die komplementären Adaptersequenzen ([www.illumina.com](http://www.illumina.com)).

Die Cluster Erzeugung beginnt mit der zufälligen Hybridisierung eines *forward* Strangs der Exom-Bibliothek mit der *Flowcell-Adaptersequenz* (Abbildung 13).

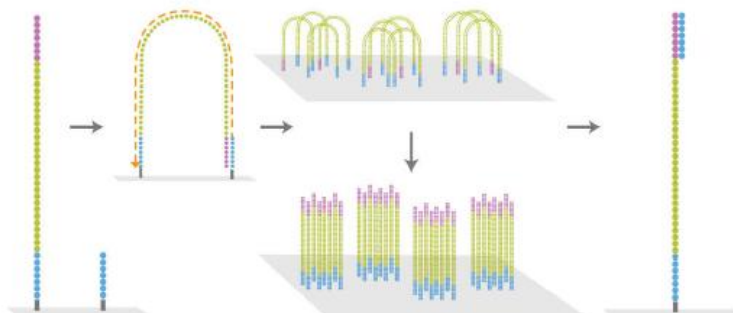
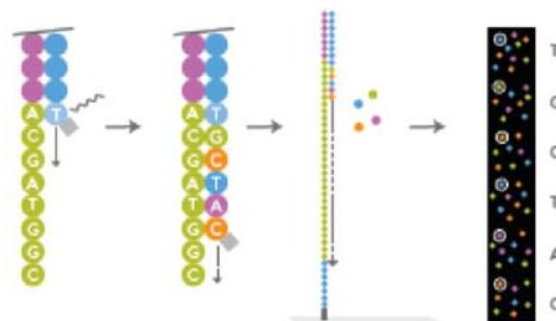


Abbildung 13: Erzeugung von Cluster. Der kovalent gebundene *forward* Strang bildet eine Brücke und wird amplifiziert. Nach Linearisierung wird der Vorgang mehrmals wiederholt. Nach Abspaltung aller *reverse* Stränge auf der *flowcell*, liegen Cluster bestehend aus *forward* Strängen vor. Anschließend kann die Anlagerung des Sequenzierprimers erfolgen. Abbildung modifiziert nach (Ansorge 2009).

Der DNA-Strang wird durch eine Polymerase in 3`-Richtung amplifiziert. Danach wird der Originalstrang der DNA entfernt, so dass nur ein kovalent an die *flowcell* gebundener *forward* Strang übrig bleibt. Anschließend hybridisiert das freie Ende des *forward* Strangs erneut. Durch eine sogenannte Brücken-Amplifikation entsteht ein doppelsträngiges DNA-Molekül. Nach Denaturierung liegen beide Stränge kovalent gebunden vor. Dieser Vorgang wird mehrmals wiederholt, so dass ein Cluster entsteht (Abbildung 13). Abschließend wird jeder *reverse* Strang von der *flowcell* abgespalten, so dass ein Cluster nur aus *forward* Strängen besteht. Die 3`-Enden dieser Stränge und der freien Adaptersequenzen auf der *flowcell* werden blockiert, um erneute Brückenbildung zu verhindern (www.illumina.com).

### 3.5.2. Sequenzierungsplattform Illumina

Das Prinzip von Illumina basiert auf Sequenzierung durch Synthese (*Sequencing By Synthesis*, SBS). Die vier Nukleotide sind mit unterschiedlich fluoreszierenden Terminatoren markiert und befinden sich in einem Reaktionsansatz (Abbildung 14).



**Abbildung 14: Sequenzierung durch Synthese. Nach Einbau der ersten Base wird die Fluoreszenz detektiert. Anschließend wird der Terminator abgespalten und ein neues Nukleotid lagert sich an. Die Basenabfolge wird durch die Fluoreszenzsignale bestimmt. Abbildung modifiziert nach (Ansorge 2009).**

Bei der Sequenzierung kommt es zum Einbau des entsprechenden Nukleotids und zur Fluoreszenz. Nach Detektion der Fluoreszenz wird der Terminator, der den Einbau von weiteren Nukleotiden verhindert, entfernt und ein neues Nukleotid kann eingebaut werden. Zunächst erfolgt die Sequenzierung der *forward* Stränge durch Anlagerung eines Sequenzierprimers. Nach Abschluss kommt es zu einer

Umkehrung (*turnover*), wodurch zunächst die sequenzierten Stränge entfernt werden. Dann werden die 3`-Enden der Stränge und die freien Adaptersequenzen auf der *flowcell* deblockiert, wodurch erneut Brücken gebildet werden. Nach Amplifikation wird der Doppelstrang denaturiert und alle *forward* Stränge von der *flowcell* abgespalten. Nun kommt es zur Anlagerung des *reverse* Primers und zur Sequenzierung ([www.illumina.com](http://www.illumina.com)).

### 3.5.3. Cluster Erzeugung und Sequenzierung Genome Analyzer Iix

Für die Erzeugung der Cluster wurde der cBot von Illumina mit dem Paired-End Genomic DNA Sample Prep Kit (Illumina) verwendet. Die Durchführung erfolgte nach Angaben des Herstellers, wobei auf eine Kapillare der *flowcell* jeweils eine Bibliothek gegeben wurde. Die Sequenzierung der in Cluster vorliegenden Exom-Bibliotheken erfolgte mit dem Genome Analyzer Iix, wobei *forward* und *reverse* für je 54 bp sequenziert wurde. Mit einer Sequenzlänge von 2 x 50 bp werden bei dem Genome Analyser Iix pro Lauf und pro *flowcell* 25 Gb – 30 Gb Sequenzen generiert. Maximal können 95 Gb Sequenzen generiert werden.

### 3.5.4. Cluster Erzeugung und Sequenzierung HiSeq2000

Mit dem cBot und dem TruSeq™ Paired end Cluster Kit v.2.5 (Illumina) wurden die Cluster für die Sequenzierung erzeugt. Die Durchführung erfolgte nach Angaben des Herstellers. Wenn die Exom-Bibliotheken 3-fach vermischt vorlagen, wurde pro Bibliothek eine Kapillare der *flowcell* verwendet. Bei 6-fach vermischten Bibliotheken wurden pro Bibliothek zwei Kapillaren verwendet. Die Sequenzierung der Exome erfolgte mit dem HiSeq2000. Es wurde der TruSeq™ SBS-Kit HS (Illumina) nach Angaben des Herstellers verwendet. Die DNA-Cluster wurden *forward* und *reverse* für je 100 bp sequenziert. Mit einer Sequenzlänge von 2 x 100 bp werden bei dem HiSeq2000 pro Lauf und pro *flowcell* 270 Gb – 300 Gb (zwei *flowcells*: 540 Gb – 600 Gb) Sequenzen generiert. Maximal können 600 Gb Sequenzen generiert werden.

---

## 3.6. Auswertung der Exom-Sequenzen

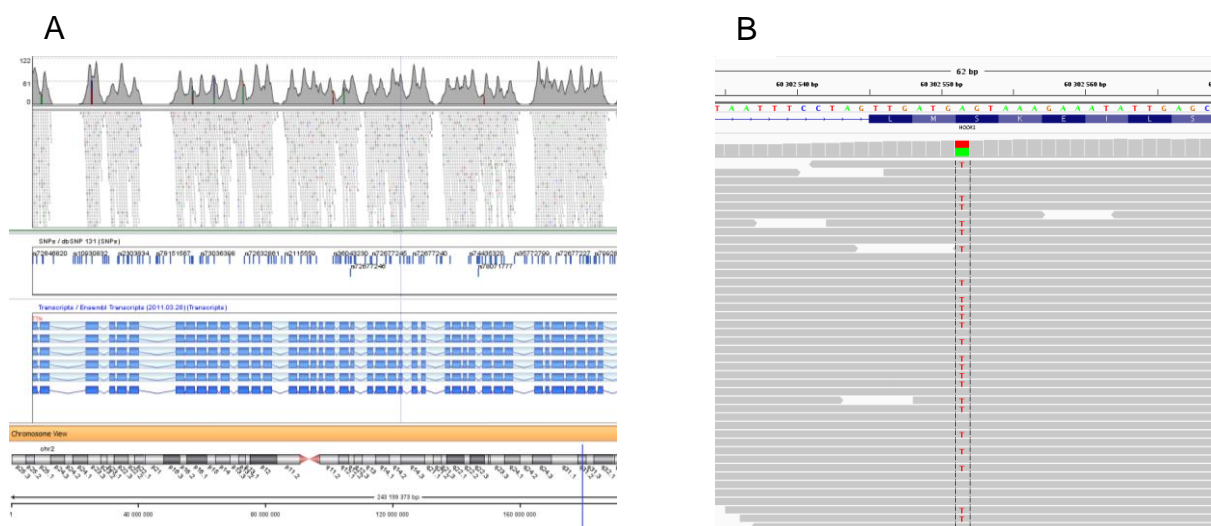
### 3.6.1. Erwartungsbereich der Information für generierte Exom-Sequenzen

Durch die Sequenzierung des menschlichen Genoms im Jahr 2001 wurden 700 - 1000 neue Varianten pro Exom diskutiert (Venter *et al.*, 2001). In der Arbeit von Sarah Ng *et al.* wurden Exom-Sequenzen generiert und der Referenzsequenz zugeordnet. Durchschnittlich wurden pro Exom 6 Gigabasen Sequenz der Referenz zugeordnet, wobei 49% aller Sequenzen im Bereich der Sonden lagen, die das Exom abdeckten (*on target*). Um Varianten im Exom eindeutig identifizieren zu können, ist eine *coverage* von 8-fach oder höher wünschenswert. 96,3% aller Basen erreichten in der Arbeit von Sarah Ng *et al.* eine *coverage* von 8-fach (Ng *et al.*, 2009). Durchschnittlich wurden 17.272 kodierende Varianten pro Individuum identifiziert, von denen 92% bereits in öffentlichen Datenbanken beschrieben wurden (Ng *et al.*, 2009). In einer weiteren Arbeit von Sarah Ng *et al.* wurde die Anzahl von neuen Varianten pro Exom kalkuliert. Ungefähr 500 - 600 neue nicht-synonyme Varianten, Spleiß-Varianten und Insertionen/Deletionen wurden pro Exom identifiziert (Ng *et al.*, 2010b).



### 3.6.2. Visualisierung mit dem Integrative Genomics Viewer (IGV)

Mit dem Integrative Genomics Viewer (Broad Institute, Cambridge, USA) konnte das *Alignment* einer Sequenz an das Genom visualisiert werden (Robinson *et al.*, 2011). Dadurch wurde die Qualität des *capturing* und der *coverage* ersichtlich (Abbildung 15A). Außerdem konnte die *Alignment*-Qualität einer Variante neben den Qualitätsangaben von SAMtools (<http://samtools.sourceforge.net/>) nochmals überprüft werden (Abbildung 15B). Bei Varianten, die sich am Ende von *reads* befinden, handelt es sich oftmals um falsch positive Varianten, da die Fehlerrate in der Sequenz gegen Ende der *reads* steigt (Day-Williams & Zeggini 2011). Abbildung 15B zeigt einen Bereich im Gen *HOOK1*, der sehr gut abgedeckt wurde. Der Basenaustausch der Variante, die detektiert wurde, wurde oft identifiziert und befindet sich hauptsächlich mittig der *reads*.



**Abbildung 15: Visualisierung der Exom-Sequenzen mit IGV. A: Qualität von *coverage* und *capturing* einer Genregion (*TTN*). B: *Alignment*-Qualität einer heterozygoten Variante im Gen *HOOK1*.**

### 3.6.3. Bioinformatische Auswertungsstrategie 1 und 2

In Tabelle 5 ist die bioinformatische Auswertungsstrategie 1 aufgeführt. Die erhaltenen Sequenzen wurden mit Burrows-Wheeler-Aligner (BWA v.0.5.8) an das menschliche Genom (NCBI36/hg18) *aligned* (<http://bio-bwa.sourceforge.net/>). In allen Exomen wurden Einzelnukleotidvarianten (*single nucleotide variants*, SNVs) und Insertionen/Deletionen (*insertion/deletions*, Indels) mittels SAMtools v.0.1.7 identifiziert (<http://samtools.sourceforge.net/>). Von diesen Varianten wurden alle Polymorphismen ausgeschlossen, die in der Datenbank dbSNP132 mit einer *minor allele frequency* (MAF) >0,5% annotiert sind. Für diese Filterung der Varianten wurde der varFilter von SAMtools verwendet, allerdings wurde die maximale *coverage* auf 9999 (-D [9999]) gesetzt (vgl. 8.4, Tabelle 51). Nachträglich wurden außerdem alle Varianten ausgeschlossen, die bereits in 600 Kontroll-Exomen (Patienten mit nicht verwandten Krankheiten und gesunde Individuen) identifiziert wurden. Die Kontroll-Exome wurden am Helmholtz-Zentrum München (Neuherberg, Deutschland) sequenziert.

**Tabelle 5: Vergleich der bioinformatischen Auswertung**

	<b>Auswertungsstrategie 1</b>	<b>Auswertungsstrategie 2</b>
<b>Sequenzierungsplattform</b>	Illumina Genome Analyzer IIx	Illumina HiSeq2000
<b>Alignment-Software</b>	Burrows-Wheeler-Aligner (BWA)	CASAVA v.1.8 (Illumina)
<b>Zusätzlich zum Alignment</b>	- Entfernen von Duplikaten	Zuordnung der Indices, Entfernen von Duplikaten
<b>Länge der reads für das Alignment</b>	54 bp (Gesamtlänge 54 bp)	75 bp (Gesamtlänge 100 bp)
<b>Referenzgenom</b>	<i>Human</i> (NCBI36/hg18)	<i>Human</i> (GRCh37/hg19)
<b>Identifizierung von Varianten (SNVs und Indels), Filterung</b>	SAMtools v.0.1.7 varFilter	SAMtools v.0.1.16
<b>Zusätzliche Annotation</b>	-	ANNOVAR
<b>Ausschluss von Varianten, Datengrundlage</b>	dbSNP132, MAF >0,5%; 600 Kontroll-Exome	dbSNP132

Bei Auswertungsstrategie 2 wurden die erhaltenen Sequenzen mit der Illumina Software CASAVA v.1.8 an das menschliche Genom (GRCh37/hg19) *aligned* (Tabelle 5). In allen Exomen wurden SNVs und Indels mittels SAMtools v.0.1.16 identifiziert. Die Filterung der Varianten nach dbSNP132 wurde ebenfalls mit SAMtools durchgeführt (vgl. 8.4, Tabelle 51).

Mit der Software ANNOVAR wurden anschließend die Regionen der Varianten in verschiedene Kategorien (Tabelle 6) untergliedert und somit genau annotiert ([www.openbioinformatics.org/annovar](http://www.openbioinformatics.org/annovar)).

**Tabelle 6: Annotationskategorien bei ANNOVAR**

<b>Kodierende Regionen</b>	Exonisch Nicht-kodierende RNA exonisch Spleiß-Varianten (Variante innerhalb einer Spleißstelle)
<b>Nicht-kodierende Regionen</b>	Intronisch Nicht-kodierende RNA intronisch Spleiß-Varianten (Variante innerhalb einer Spleißstelle) Varianten zwischen Genen 3'-UTR 5'-UTR <i>Downstream</i> (Varianten innerhalb 1 kbp von Transkriptionseende) <i>Upstream</i> (Variante innerhalb 1 kbp von Transkriptionsstart)

### 3.6.4. Zusätzliche Qualitätsfilter für beide Strategien

Während der Detektion von Varianten wurden verschiedene Parameter durch die entsprechende Software angegeben (Ewing *et al.*, 1998; Li *et al.*, 2009; Danecek *et al.*, 2011). Diese Parameter wurden genutzt, um bei der Auswertung der Exom-Sequenzen zusätzliche Qualitätswerte für eine Variante festzulegen und danach zu filtern (Tabelle 7). Für die Auswertungsstrategie 1 sollte die Qualität einer Variante größer 15 (Qual >15) sein. Außerdem sollte die Variante in 15% aller *reads* detektiert werden und davon in mindestens 5% in unterschiedlichen *reads*, *forward* oder *reverse* vorkommen. Für die Auswertungsstrategie 2 sollte die Qualität einer Variante größer 20 (Qual >20) sein und die Variante mindestens dreimal abgebildet werden (DP >3). Zudem wurde die *mapping quality* und *genotype quality* auf größer 20 (MQ und GQ >20) gesetzt. Eine Variante musste mindestens einmal im *forward* Strang detektiert werden, wenn in den *reverse* Strängen die Variante zehnmal oder öfter detektiert wurde und umgekehrt. Da heterozygote Varianten bevorzugt wurden, sollte die *consensus quality* FQ positiv sein. Zuletzt wurden nur Varianten ausgewählt, die mit mindestens 80% aller *reads* ausgewertet wurden (Tabelle 7).

Die Kriterien dieser Parameter sind zusammenfassend in Tabelle 7 aufgeführt und die Parameter in der Legende zu Tabelle 7 erklärt.

**Tabelle 7: Vergleich der zusätzlichen Qualitätsfilter**

	<b>Auswertungsstrategie 1</b>	<b>Auswertungsstrategie 2</b>
<b>Zusätzliche Qualitätsfilter</b> <b>Parameter von SAMtools</b>	<ul style="list-style-type: none"> <li>• Qual &gt;15</li> <li>• Verhältnis Referenzallel/Nicht-Referenzallel: 0,15</li> <li>• Verhältnis <i>forward/reverse read</i> mit SNV: 0,05</li> </ul>	<ul style="list-style-type: none"> <li>• Qual &gt;20</li> <li>• DP &gt;3</li> <li>• MQ &gt;20</li> <li>• GQ &gt;20</li> <li>• SNV in <i>forward</i> oder <i>reverse read</i> &gt; 1, wenn in Gegenstränge SNV <math>\geq</math> 10 vorliegt</li> <li>• FQ bei heterozygot positiv</li> <li>• DPused/DP &gt; 0,8</li> </ul>

Legende Tabelle 7: MQ: *mapping quality*; angegeben mit dem Phred score; MQ >20 bedeutet, dass ein *read* von 100 falsch zugeordnet wurde.

GQ: *genotype quality*; angegeben mit dem Phred score; GQ >20 bedeutet, dass ein Genotyp von 100 falsch angegeben wurde.

DP: *raw read depth*; Gesamtanzahl an *reads* bei einer bestimmten Basenposition (ohne Qualitätsfilter); entspricht der *coverage*.

DP\_used: Anzahl an *reads*, die verwendet wurden, d.h. die die Qualitätsfilter passiert haben.

DP\_1: Anzahl der *forward reads*, bei denen das Referenzallel detektiert wurde.

DP\_2: Anzahl der *reverse reads*, bei denen das Referenzallel detektiert wurde.

DP\_3: Anzahl der *forward reads*, bei denen das mutierte Allel detektiert wurde.

DP\_4: Anzahl der *reverse reads*, bei denen das mutierte Allel detektiert wurde.

Qual: *variant/reference quality*; angegeben mit dem Phred score; Qual >20 bedeutet, dass eine Variante von 100 falsch angegeben wurde.

FQ: *consensus quality*; angegeben mit dem Phred score; bei verschiedenen Genotyp = heterozygot → FQ positiv, bei gleichen Genotyp = homozygot → FQ negativ.

### 3.6.5. Filtern von paralogen Genen

Bei der Auswertung der Exom-Sequenzen konnte ein zusätzlicher Filter geschaltet werden, um paraloge Gene bzw. Genbereiche zu identifizieren. Von den Varianten wurde die Sequenz +/- 100 bp von der Austauschposition in dem *Tool* Blat des UCSC Genome Browser analysiert ([www.genome.ucsc.edu](http://www.genome.ucsc.edu)). Trat bei der Gesamtlänge der eingegebenen Sequenz eine Paralogie von >99% zu einer anderen Sequenz auf, wurde der Genbereich als paralog eingestuft.

### 3.6.6. Vorhersagemodul PolyPhen-2

Die identifizierten Varianten der Exom-Sequenzierungen wurden mit der Software PolyPhen-2 als schädlich oder nicht schädlich für das Protein eingestuft. PolyPhen-2 ist eine bioinformatische Plattform, um den Einfluss eines Aminosäureaustausches auf die Funktion und Struktur eines entstehenden Proteins vorherzusagen. Zur Vorhersage werden die Informationen von Proteinstrukturen von PDB/DSSP sowie die Proteinsequenzen von UniProtKB/UniRef100 verwendet (<http://genetics.bwh.harvard.edu/pph2/>).

## 3.7. Mutationsscreening, Genotypisierung und Expression

### 3.7.1. IDAHO-Screening

Mit dem LightScanner® System (Idaho Technology Inc., Salt Lake City, USA) kann mittels *high resolution melting* (HRM) ein Mutationsscreening durchgeführt werden. Ein spezieller interkalierender Farbstoff, LCGreen Dye, sorgt für eine gute Detektion von Sequenzvarianten. Die zu untersuchenden Amplikons werden mit dem LCGreen Dye im PCR-Ansatz erzeugt. Anschließend erfolgt eine Aufschmelzung der PCR-Produkte mit dem LightScanner®. Die Fluoreszenz nimmt im Verlauf der Aufschmelzung immer weiter ab. PCR-Produkte mit unterschiedlicher Basenfolge weisen dabei einen unterschiedlichen Kurvenverlauf auf. Es können nicht nur SNPs und SNVs sondern auch Insertionen und Deletionen detektiert werden (Vossen *et al.*, 2009).

### Vorbereiten der DNA

Die Konzentration der DNA in den 96 *well*-Platten (Mutterplatten) wurde auf 10 ng/μl normalisiert. Ausgehend davon wurde die DNA auf eine Konzentration von 1 ng/μl mit H<sub>2</sub>O verdünnt. Anschließend wurden 5 μl der DNA mit dem TeMo Roboter (Tecan, Männedorf, Schweiz) auf 384 *well*-Platten verteilt und über Nacht getrocknet. Von jeder DNA Probe wurden Duplikate vorgelegt.

### Amplifikation mit LCGreen Dye

Zu der getrockneten DNA wurden 5 µl Mastermix gegeben. Der Mastermix enthielt 3 µl H<sub>2</sub>O, 0,5 µl 10x Puffer, 0,5 µl dNTPs (2 mM), 0,5 µl MgCl<sub>2</sub> (25 mM), 0,25 µl LCGreen Dye Plus (Idaho Technology Inc.), 0,05 µl Taq-Polymerase und je 0,1 µl *forward* und *reverse* Primer (10 µM). Amplifiziert wurde mit einem speziellen Touchdown-PCR Programm:

1. Denaturierung 95°C	10 min	
2. Denaturierung 95°C	30 sec	
3. <i>Annealing</i> 70°C	30 sec	-0,5°C pro Zyklus
4. Elongation 72°C	30 sec	

Die Schritte zwei bis vier wurden 25-mal zyklisch wiederholt.

5. Denaturierung 95°C	30 sec
6. <i>Annealing</i> 54°C	30 sec
7. Elongation 72°C	30 sec

Die Schritte fünf bis sieben wurden ebenso 25-mal zyklisch wiederholt.

8. Denaturierung 94°C	30 sec
9. finale Elongation 72°C	10 min
10. Abkühlung 25°C	30 sec

### Schmelzkurve aufnehmen mit LightScanner® System und Auswertung

Zur Aufnahme der Schmelzkurve wurden die PCR-Produkte mit 8 µl Öl überschichtet. Die Einstellungen des LightScanners waren wie folgt:

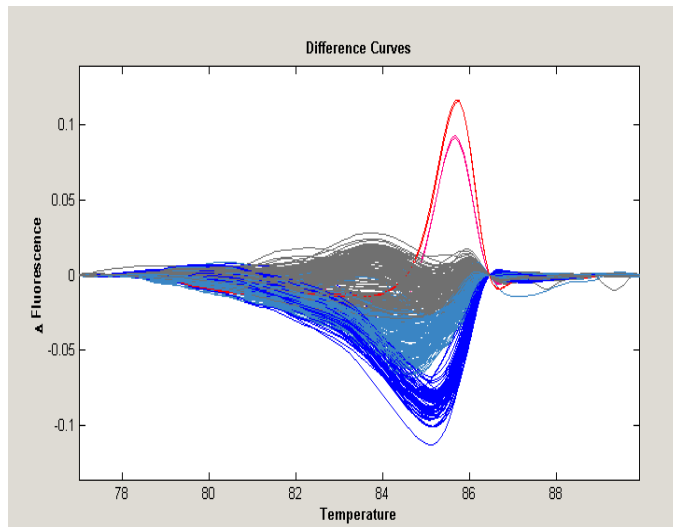
*Hold* Temperatur: 74°C

*Start* Temperatur: 77°C

*End* Temperatur: 96°C

Die Auswertung der Schmelzkurven der Amplifikationsprodukte erfolgte mit der LightScanner™ Software v.1.1.0.566 (Idaho Technology Inc.). Die Referenzsequenz wurde als größte Gruppe von gleichverlaufenden Schmelzkurven definiert und ein Kurvenverlauf festgelegt. Bei einer auffälligen Abweichung des Kurvenverlaufs von der Schmelzkurve der Referenzsequenz wurden die betreffenden PCR-Produkte sequenziert (vgl. IDAHO-Sequenzierung). Dabei war entscheidend, dass die Duplikate auf der Platte die gleiche Abweichung zeigten. Exemplarisch ist die Auswertung von 384 Schmelzkurven (entspricht 192 PCR-Produkte im Duplikat) in Abbildung 16 dargestellt. Man konnte fünf verschiedene Gruppen von

Schmelzkurven festlegen. Von den grau hinterlegten Schmelzkurven der Referenzsequenz wichen die hellblau und dunkelblau markierten Schmelzkurven deutlich ab. Durch Sequenzierung einzelner PCR-Produkte aus den Gruppen, konnten jeweils SNPs die Abweichung erklären. Außerdem zeigten die PCR-Produkte zweier Duplikate eine starke Abweichung des Kurvenverlaufs, dargestellt in roten und rosa Kurven. Die Sequenzierung der Produkte führte zur Identifizierung zweier neuer SNVs (Abbildung 16).



Legende Abbildung 16:

Grau: Schmelzkurve der Referenzsequenz

rot: Schmelzkurve mit SNV C610Y

rosa: Schmelzkurve mit SNV V587I

hellblau: Schmelzkurve mit SNP rs2306554

dunkelblau: Schmelzkurve mit SNP rs2306555

**Abbildung 16: Auswertung der Schmelzkurven eines IDAHO-Screenings. Auf der Abszisse ist die Temperatur aufgetragen, auf der Ordinate die gemessene Fluoreszenz.**

### IDAHO-Sequenzierung

Zur Aufreinigung wurden zu 8 µl der PCR-Produkte 100 µl H<sub>2</sub>O gegeben und das Gemisch in eine Millipore-Patte überführt. Anschließend wurde mit einem Vakuumblock bei 20 inches Hg für 8 min die Flüssigkeit abgesaugt. Um die Nukleinsäuren von der Membranoberfläche zu lösen, wurden 20 µl H<sub>2</sub>O zugegeben, gemischt und die aufgereinigten PCR-Produkte in eine neue Platte überführt.

Für die Sequenzierungsreaktion wurden 1 µl PCR-Produkt, 1 µl *forward* bzw. *reverse* Primer (10 µM), 0,5 µl BIG DYE (Applied Biosystems), 1,5 µl BIG DYE-Puffer und 1 µl H<sub>2</sub>O zusammen pipettiert und mit folgendem Programm behandelt:

1. Denaturierung 94°C	30 sec
2. <i>Annealing</i> 50°C	15 sec
3. Elongation 60°C	4 min
4. Abkühlung 4°C	1 min

Die Schritte eins bis drei wurden 24-mal zyklisch wiederholt.

Die Sequenzierungsprodukte wurden mit 25 µl 100% EtOH und anschließend mit 125 µl 70% EtOH gefällt. Nach Zugabe von 25 µl H<sub>2</sub>O wurde auf einem 3730 DNA Analyzer (Applied Biosystems) sequenziert. Die Auswertung der Sequenzen erfolgte mit SeqScape v.2.7 (Applied Biosystems).

### 3.7.2. Genotypisierung mit KASPar Assay

Für die Genotypisierung einer Variante wurde der Genotypisierungs-Assay von KBioscience (Hoddesdon, England) verwendet. Das Prinzip basiert auf dem Fluoreszenz-Resonanz-Energie-Transfer (FRET). Im Reaktionsansatz befinden sich zwei mit einem unterschiedlichen Fluorophor-markierte Oligonukleotide. In freier Lösung hybridisieren diese mit Oligonukleotide, die mit einem Quencher markiert sind. Dadurch wird die Emission von Licht verhindert. Für eine Genotypisierung sind ein nicht-markierter *reverse* Primer und ein Allel-spezifischer *forward* Primer im Reaktionsansatz von Wichtigkeit. Die Allel-spezifischen *forward* Primer sind jeweils um die komplementäre Sequenz der Fluorophor-markierten Oligonukleotide verlängert. Bei Ablauf der PCR kommt es zu einer Allel-spezifischen Amplifikation der DNA-Stränge, die dadurch um die komplementäre Sequenz der Fluorophor-markierten Oligonukleotide verlängert werden. Da die Schmelztemperatur von einem Fluorophor-markiertem Oligonukleotid kleiner ist als die *annealing* Temperatur der PCR, kommt es zu der Auflösung der Fluorophor-Quencher Paare und zu einer Anlagerung an die DNA-Stränge. Dadurch kommt es zur Fluoreszenz (Robinson 2009).



Zum Überprüfen einer Variante im Titin-Gen (genomische Position: 179411847) wurde ein KASP Assay by Design (Tabelle 8) bestellt und der Versuch nach Angaben des Herstellers (KBioscience) durchgeführt. Das Referenzallel wurde mit dem Fluorophor VIC, das mutierte Allel mit dem Fluorophor FAM detektiert. Die Detektion der Fluoreszenz und die anschließende Auswertung erfolgte mit dem ABI PRISM® 7000 Sequence Detection System (Applied Biosystems) und mit der Software SDS v.2.3 (Applied Biosystems).

**Tabelle 8: KASP Assay by Design für eine Titin-Variante**

Assay	Allel FAM	Allel VIC	CG%_FAM	CG%_VIC
<i>TTN_179411847</i>	A	T	45.8	45.8

### 3.7.3. Quantitative PCR

Die Expression von Genen kann mit Hilfe der quantitativen PCR, auch *real-time* PCR oder qPCR genannt, untersucht werden. Ausgangsmaterial stellte cDNA von humanen Kardiomyozyten (Human Cardiac Myocytes; PromoCell, Heidelberg, Deutschland) und von verschiedenen Geweben dar (Human MTC Panel I; Clontech, Mountain View, USA). Die Konzentration der cDNA betrug 5 ng pro Reaktion. Verwendet wurden Genexpressions-Assays (Tabelle 9) nach Angaben des Herstellers (Applied Biosystems). Diese enthalten nicht-markierte Oligonukleotide zur Amplifikation der cDNA und eine TaqMan-Sonde, welche innerhalb der Sequenz mit der cDNA hybridisiert. Die Taqman-Sonde trägt am 5`-Ende einen Fluorophor (z.B. FAM) und am 3`-Ende einen nicht fluoreszierenden Quencher. Durch FRET wird die Emission von Licht verhindert. Bei der Elongation wird durch die 5`-Exonuklease-Aktivität der Polymerase die TaqMan-Sonde abgebaut, wodurch der Fluorophor freigesetzt wird (Lazaruk *et al.*, 2006). Die auftretende Fluoreszenz wurde mit dem ABI PRISM® 7900HT Sequence Detection System (Applied Biosystems) detektiert und mit der Software SDS v.2.3 (Applied Biosystems) ausgewertet.

**Tabelle 9: Genexpressions-Assays**

Assay ID	Fluorophor
<i>GJB4</i> -Hs00920816_s1	FAM
<i>ZG16</i> -Hs00380609_m1	FAM
<i>INTS1</i> -Hs01357109_m1	FAM
<i>HOOK1</i> -Hs00210665_m1	FAM
<i>GAPDH</i> -Hs99999905_m1	FAM

### 3.7.4. Berechnung der relativen Genexpression

Für die Berechnung der relativen Expression eines Gens in Bezug auf ein Referenzgen wurde die delta-delta Ct-Methode ( $Ct = \textit{threshold cycle}$ ) verwendet (Pfaffl 2001). Als Referenzgen wurde GAPDH verwendet. Als Kalibrator wurde die Expression im Skelettmuskel festgelegt.

$$\Delta Ct \text{ Zielgen} = Ct \text{ Zielgen} - Ct \text{ Referenzgen}$$

$$\Delta\Delta Ct = \Delta Ct \text{ Zielgen} - \Delta Ct \text{ Kalibrator}$$

$$\text{Relative Expression} = 2^{-\Delta\Delta Ct}$$

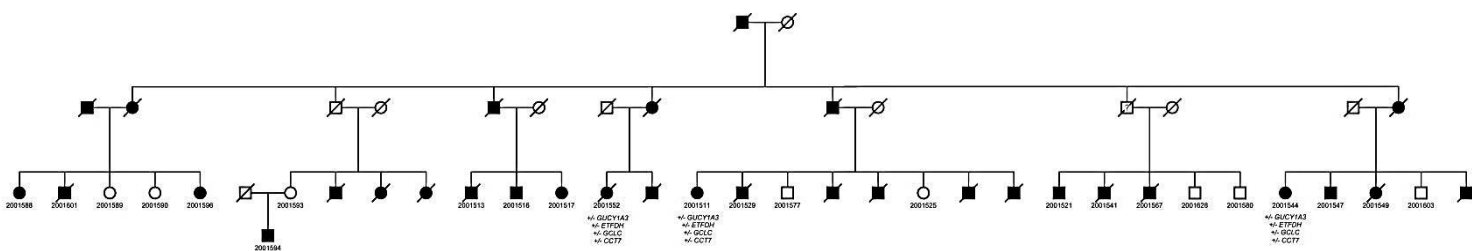
## 4. Ergebnisse

### 4.1. Exom-Sequenzierung Herzinfarkt-Großfamilie

Für die Anreicherung der Exom-Anteile wurde der SureSelect Human All Exon Kit von Agilent Technologies verwendet (vgl. 3.4.1 und 3.4.2). Die Sequenzierung erfolgte mit dem Genome Analyzer IIx (vgl. 3.5.1 - 3.5.3). Die Auswertung der Sequenzen wurde mit der Auswertungsstrategie 1 durchgeführt (vgl. 3.6.3 und 3.6.4).

#### 4.1.1. Exom-Sequenzierung HI-Großfamilie 2001

In Abbildung 17 ist der Stammbaum der HI-Großfamilie 2001 gezeigt. Eine vergrößerte Darstellung findet sich in Abbildung 40 unter 8.5. 31 Individuen dieser Familie erlitten einen Herzinfarkt oder waren an der koronaren Herzerkrankung erkrankt.



**Abbildung 17: Stammbaum HI-Großfamilie 2001. Quadrat: Mann, Kreis: Frau, ausgefüllte Symbole: betroffen, offene Symbole: nicht betroffen, durchgestrichene Symbole: verstorben, ?: Phänotyp unbekannt.**

Von 15 betroffenen Individuen war DNA vorhanden. Drei Cousins (ID 2001552, ID 2001511, ID 2001544), die im Alter zwischen 47 und 53 Jahren einen Herzinfarkt erlitten, wurden für die Exom-Sequenzierung in dieser Familie ausgewählt. Die klinischen Charakteristika der ausgewählten Familienmitglieder sowie der weiteren Familienmitglieder sind in Tabelle 44 unter 8.1 aufgeführt.

Pro Exom wurden durchschnittlich 119 Millionen *reads* (*forward* und *reverse*) erzeugt (Tabelle 10). Da Duplikate von *reads* durch die Amplifikationsschritte im Protokoll entstehen, wurden 3,9% – 4,7% aller *reads* entfernt. Als Duplikate wurden *reads* bezeichnet, die identische Länge aufweisen und die exakt an der gleichen Position der Referenzsequenz zugeordnet wurden. Ungefähr 96% der verbleibenden *reads*

(110.562.838 - 117.477.400 *mapped reads*) wurden der Referenzsequenz zugeordnet, wovon circa 53% im Bereich der Sonden (*on target*) lagen. Die insgesamt etwa 6 Gigabasen (Gb) pro Ansatz erzeugte Sequenz war gleichmäßig über das Exom verteilt, d.h. 99% aller Basen waren zumindest einmal abgebildet (*coverage*  $\geq 1x$ ). Eine *coverage* von  $\geq 20x$  wurde angestrebt, um heterozygote Varianten mit Sicherheit detektieren zu können. Diese *coverage* lag bei 85% aller Basen vor (Tabelle 10).

**Tabelle 10: Information der generierten Exom-Sequenzen für HI-Großfamilie 2001**

ID	Duplikate in %	<i>reads/ mapped reads</i>	<i>mapped reads</i> in %	<i>mapped sequence</i> in bp	<i>on target</i> in %	<i>coverage</i> $\geq 1x$ in %	<i>coverage</i> $\geq 4x$ in %	<i>coverage</i> $\geq 8x$ in %	<i>coverage</i> $\geq 20x$ in %
2001552	4,0	115.565.636/ 110.562.838	95,7	5.970.393.252	51	99	98	95	84
2001511	3,9	121.973.552/ 117.477.400	96,3	6.343.779.600	53	99	97	95	85
2001544	4,7	119.951.616/ 114.964.538	95,8	6.208.085.052	55	99	98	95	86

Nach dem Prozess des *mappings* wurden die Varianten mit SAMtools identifiziert. In Tabelle 11 ist die Anzahl der detektierten Varianten der einzelnen Exome aufgelistet. Dabei wurde zwischen homozygoten und heterozygoten Varianten unterschieden. Durch das *capturing* wurden auch Varianten detektiert, die in intronischen Bereichen liegen. Für ID 2001552 wurden z.B. 23.585 Varianten *gecaptured*, von denen nur 16.320 in kodierenden Bereichen liegen. Da ein autosomal-dominantes Vererbungsmuster vermutet wurde, wurden heterozygote, nicht-synonyme Varianten, Spleiß-Varianten und Insertionen/Deletionen ohne rs-Nummer bevorzugt. Bei 12% - 13% aller ausgewählten Varianten lag kein Eintrag in dbSNP132 vor. Bei ID 2001552 wurden 602 Varianten ausgewählt. Für ID 2001511 waren es 576 ausgewählte Varianten und für ID 2001544 kamen 609 ausgewählte Varianten in Frage (Tabelle 11).

**Tabelle 11: Anzahl detektierter Varianten der Exome von HI-Großfamilie 2001**

ID	Typ	<i>captured/ ohne rs-Nummer</i>	ohne rs-Nummer in %	kodierend/ ohne rs-Nummer	ohne rs-Nummer in %	synonym/ ohne rs-Nummer	ohne rs-Nummer in %	nicht-synonym, Indel, Spleiß-Variante/ ohne rs-Nummer	ohne rs-Nummer in %
2001552	Alle	23585/ 1446	6,1	16320/ 995	6,1	8777/ 376	4,3	7543/ 619	8,2
	Homozygot	9156/ 72	0,8	6185/ 37	0,6	3411/ 20	0,6	2774/ 17	0,6
	Heterozygot	14429/ 1374	9,5	10135/ 958	9,4	5366/ 356	6,6	4769/ 602	12,6

**Tabelle 11: Anzahl detektierter Varianten der Exome von HI-Großfamilie 2001**

ID	Typ	captured/ ohne rs- Nummer	ohne rs- Nummer in %	kodierend/ ohne rs- Nummer	ohne rs- Nummer in %	synonym/ ohne rs- Nummer	ohne rs- Nummer in %	nicht-synonym, Indel, Spleiß- Variante/ ohne rs-Nummer	ohne rs- Nummer in %
2001511	Alle	23150/ 1403	6,1	16085/ 937	5,8	8665/ 344	4,0	7420/ 593	8,0
	Homozygot	9010/ 69	0,8	6111/ 23	0,4	3374/ 6	0,2	2737/ 17	0,6
	Heterozygot	14140/ 1334	9,4	9974/ 914	9,2	5291/ 338	6,4	4683/ 576	12,3

ID	Typ	captured/ ohne rs- Nummer	ohne rs- Nummer in %	kodierend/ ohne rs- Nummer	ohne rs- Nummer in %	synonym/ ohne rs- Nummer	ohne rs- Nummer in %	nicht-synonym, Indel, Spleiß- Variante/ ohne rs-Nummer	ohne rs- Nummer in %
2001544	Alle	23432/ 1501	6,4	16196/ 1019	6,3	8712/ 382	4,4	7484/ 637	8,5
	Homozygot	9223/ 87	0,9	6311/ 45	0,7	3401/ 17	0,5	2910/ 28	1,0
	Heterozygot	14209/ 1414	10,0	9885/ 974	9,9	5311/ 365	6,9	4574/ 609	13,3

#### 4.1.2. Identifizierte Varianten

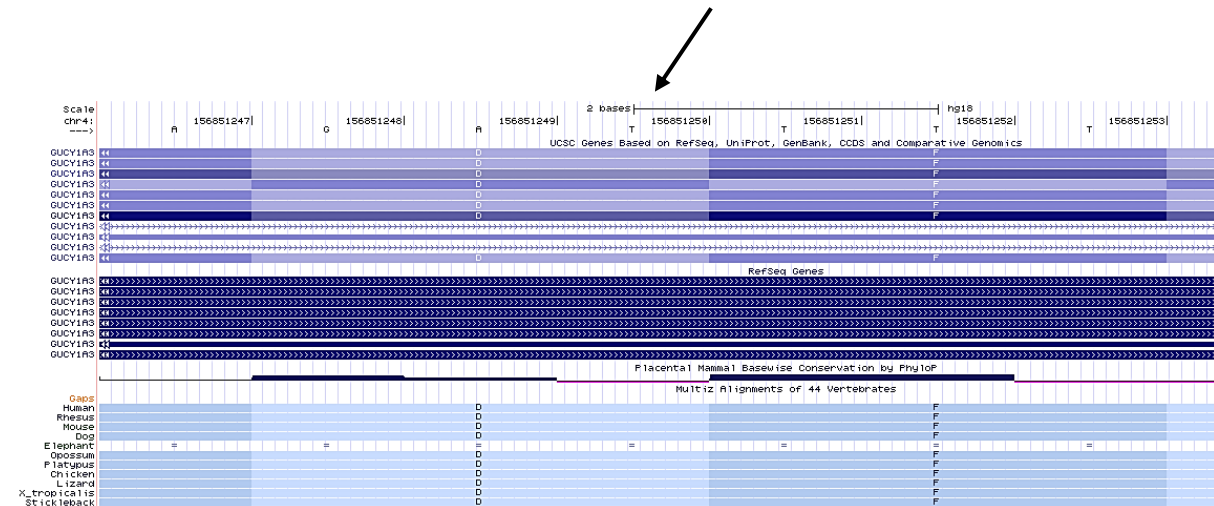
Der Vergleich der ausgewählten Varianten zwischen den Individuen ergab, dass nur vier Varianten von allen drei Betroffenen geteilt wurden. Diese Varianten sind in Tabelle 12 aufgelistet und in Abbildung 17 eingezeichnet.

**Tabelle 12: Geteilte Varianten der Exome von HI-Großfamilie 2001**

Gen	Annotation NCBI36/hg18	PolyPhen-2	Konservierung der Mutationsstelle
<i>GUCY1A3</i>	NM_000856: Position chr4:156851250: Exon 7: c.482_483insT, p.L163Ffs*24	<i>frameshift</i>	ja
<i>ETFDH</i>	NM_004453: Position chr4:159821155: Exon 2: c.C121T, p.R41X	<i>stopgain variant</i>	ja
<i>GCLC</i>	NM_001498: Position chr6:53479703: Exon 10: c.G1164T, p.H388Q	<i>benign</i>	ja
<i>CCT7</i>	NM_006429: Position chr2:73333439: Exon 12: c.C962T, p.S525L	<i>possibly damaging</i>	ja

Die Variante im Gen *GUCY1A3* führte zu einer Veränderung des Leserahmens für das entstehende Protein (*frameshift*). Eine Einführung eines Stoppkodons (*stopgain variant*) wurde durch die Variante im Gen *ETFDH* verursacht. Die Variante im Gen *GCLC* wurde über PolyPhen-2 Vorhersage als nicht schädlich für das entstehende Protein (*benign*) eingestuft, während die Variante im Gen *CCT7* als vermutlich schädlich (*possibly damaging*) eingestuft wurde (Tabelle 12).

Weiterhin wurde die Konservierung der Mutationsstelle im UCSC Genome Browser überprüft. Lag die Aminosäure, die aus dem Kodon mit der Referenzbase gebildet wird, konserviert in allen gelisteten Organismen (Wirbeltiere) vor, wurde die Mutationsstelle als konserviert angesehen. In Tabelle 12 wurde das Vorhandensein der Konservierung mit „ja“ gekennzeichnet. In Abbildung 18 ist beispielhaft die Konservierung der *GUCY1A3*-Mutationsstelle gezeigt. Es lag ein hoher Konservierungsgrad bis zum Stichling (*stickleback*) vor.



**Abbildung 18: Konservierung der Mutationsstelle der *GUCY1A3*-Insertion. Der Pfeil markiert die Position der Mutationsstelle im Gen *GUCY1A3*. Die Konservierung ist für folgende Organismen gezeigt: Mensch (*human*), Rhesusaffe (*rhesus*), Maus (*mouse*), Hund (*dog*), Opossum (*opossum*), Schnabeltier (*platypus*), Huhn (*chicken*), Eidechse (*lizard*), Frosch (*Xenopus tropicalis*) und Stichling (*stickleback*) (<http://genome-ucsc.edu/>).**

#### 4.1.3. Ko-Segregation der identifizierten Varianten

Die vier Varianten wurden ursprünglich bei den Exom-Sequenzen der Individuen ID 2001552, ID 2001511 und ID 2001544 identifiziert. Zum Überprüfen der Kausalität einer Variante wurden die Genregionen, welche die Varianten enthalten, bei allen Familienmitgliedern mittels konventioneller Sanger-Sequenzierung untersucht (Tabelle 13). Für die Auswertung der Ko-Segregation war die Einteilung in betroffen und nicht betroffen entscheidend. Alle Individuen, die einen Herzinfarkt erlitten hatten oder die Diagnose einer koronaren Herzerkrankung aufwiesen, wurden als betroffen eingestuft. Keine Variante ko-segregierte vollständig mit dem Phänotyp KHK/HI. Die *GUCY1A3*-Insertion und die *ETFDH*-Variante wurden gekoppelt vererbt und zeigten somit das gleiche Ko-Segregationsmuster. Sieben von den 15 betroffenen Familienmitgliedern und zwei von den acht nicht betroffenen Familienmitgliedern

wiesen beide Varianten auf. Die *CCT7*-Variante zeigte die beste Ko-Segregation mit dem Phänotyp, da elf der 15 Betroffenen die Variante in dem Gen hatten. Jedoch konnte aufgrund der fehlenden vollständigen Ko-Segregation, keine Variante als kausal für die Entstehung von KHK/MI betrachtet werden. Auffallend ist, dass sämtliche *GUCY1A3/CCT7*-Doppelmutationsträger (n=7) erkrankt sind (Tabelle 13).

**Tabelle 13: Ko-Segregation der identifizierten Varianten in der Familie 2001**

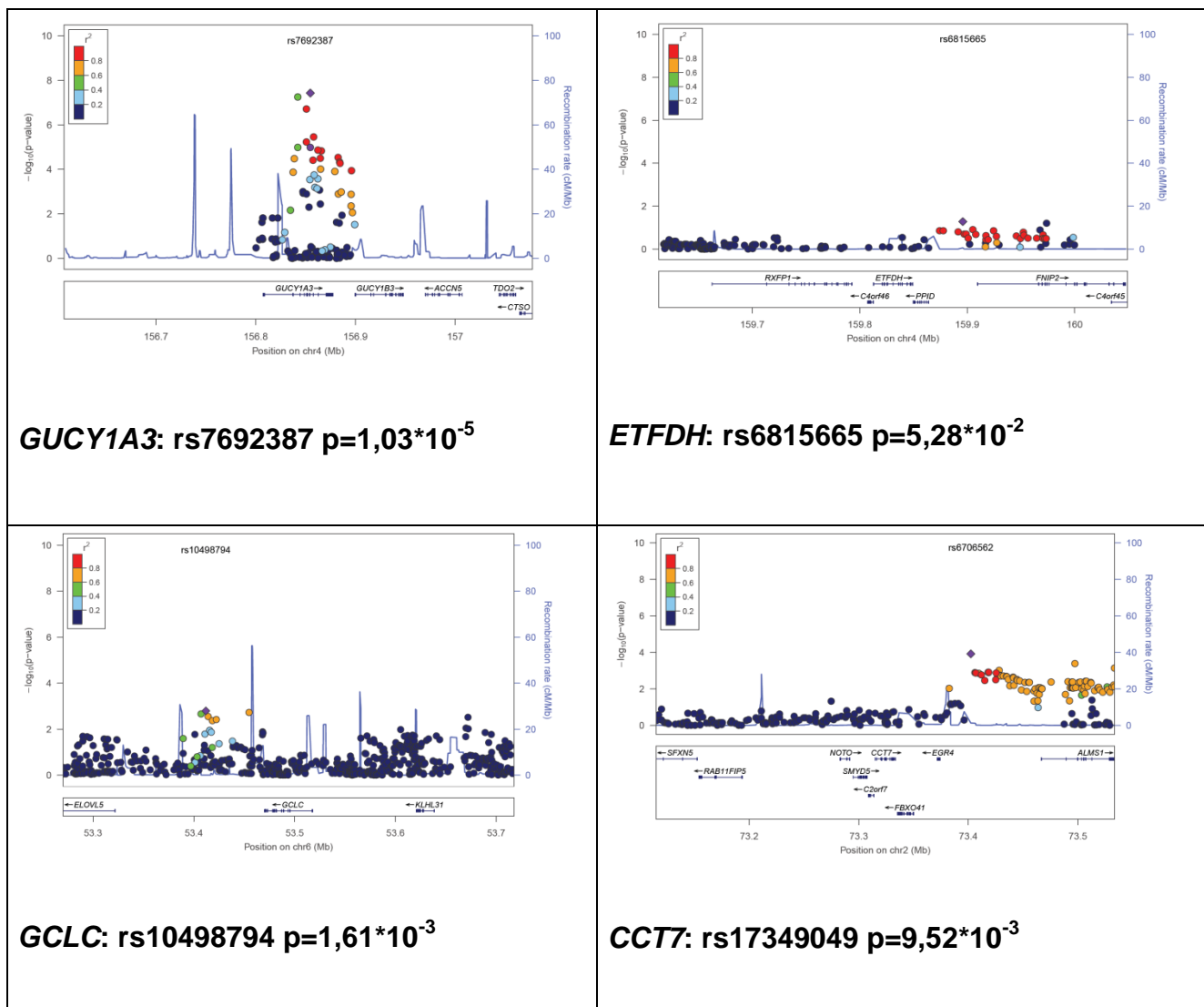
ID	Phänotyp *	<i>GUCY1A3</i> -Insertion	<i>ETFDH</i> -Variante	<i>GCLC</i> -Variante	<i>CCT7</i> -Variante
2001588	MI, 58 Jahre	positiv	positiv	negativ	positiv
2001601	KHK, 55 Jahre	positiv	positiv	negativ	positiv
2001589	Nicht betroffen, 62 Jahre	negativ	negativ	negativ	positiv
2001590	Nicht betroffen, 72 Jahre	negativ	negativ	negativ	negativ
2001596	MI, 55 Jahre	positiv	positiv	negativ	positiv
2001593	Nicht betroffen	negativ	negativ	negativ	negativ
2001594	MI, 33 Jahre	negativ	negativ	negativ	negativ
2001513	KHK, 59 Jahre	negativ	negativ	negativ	positiv
2001516	MI, 50 Jahre	negativ	negativ	negativ	positiv
2001517	KHK, 66 Jahre	negativ	negativ	negativ	negativ
2001552	MI, 53 Jahre	positiv	positiv	positiv	positiv
2001511	MI, 52 Jahre	positiv	positiv	positiv	positiv
2001529	MI, 57 Jahre	negativ	negativ	positiv	positiv
2001577	Nicht betroffen, 73 Jahre	positiv	positiv	negativ	negativ
2001525	Nicht betroffen, 70 Jahre	positiv	positiv	positiv	negativ
2001521	KHK, 60 Jahre	negativ	negativ	positiv	positiv
2001567	KHK, 59 Jahre	negativ	negativ	positiv	negativ
2001628	Nicht betroffen, 69 Jahre	negativ	negativ	positiv	negativ
2001580	Nicht betroffen, 80 Jahre	negativ	negativ	positiv	positiv
2001544	MI, 47 Jahre; KHK	positiv	positiv	positiv	positiv
2001547	KHK, 58 Jahre	positiv	positiv	positiv	positiv
2001549	MI, 67 Jahre	negativ	negativ	positiv	negativ
2001603	Nicht betroffen, 75 Jahre	negativ	negativ	positiv	positiv
	Betroffene mit Mutation/ Gesamtzahl Betroffene	8/15	8/15	8/15	11/15
	Nicht-Betroffene mit Mutation/ Gesamtzahl Nicht- Betroffene	2/8	2/8	4/8	3/8

\* Bei Betroffenen Diagnosealter, bei Unbetroffenen Einschlussalter

#### 4.1.4. Einblick und Daten der CARDIoGRAM-Studie

Eine mögliche Assoziation von häufigen Polymorphismen in den unter 4.1.2 identifizierten Genen in Fall-Kontroll-Studien wurde mit Daten aus der CARDIoGRAM-Studie (Preuss *et al.*, 2010; Schunkert *et al.*, 2011) untersucht. In der

CARDIoGRAM-Studie wurde eine Metaanalyse aus 14 genomweiten Assoziationsstudien mit insgesamt mehr als 22.000 KHK-Fällen und über 60.000 gesunde Kontrollen durchgeführt. Für die vier Gene, die durch die Exom-Sequenzierung als neue Kandidatengene für KHK/HI galten, wurden die Assoziationsplots der Studie betrachtet. Nur das Gen *GUCY1A3* zeigte mit einem p-Wert von  $1,03 \cdot 10^{-5}$  eine signifikante Assoziation mit KHK. Die Abbildungen und p-Werte wurden von Frau Prof. Jeanette Erdmann (Universität zu Lübeck) zur Verfügung gestellt (Abbildung 19).



**Abbildung 19: Assoziationsplots aus der CARDIoGRAM-Studie (Preuss *et al.*, 2010; Schunkert *et al.*, 2011). Auf der Abszisse ist die Position auf dem Chromosom in Mb angegeben, auf der linken Ordinate  $-\log_{10}$  des p-Wertes für eine Assoziation mit KHK und auf der rechten Ordinate die Rekombinationsrate in cM/Mb. Die am stärksten assoziierten SNPs sind mit den resultierenden p-Werten angegeben. Farbkodierung: Kopplungsungleichgewicht in  $r^2$  laut Legende (von rot: perfektes Kopplungsungleichgewicht bis blau: nicht gekoppelt).**



#### 4.1.5. IDAHO-Screening für ein Kandidatengen

Durch die gezeigte Assoziation mit KHK und der Funktion von *GUCY1A3* im kardiovaskulären System (siehe 5.3), stellte dieses Gen ein plausibles Kandidatengen für KHK/MI dar. Aus diesem Grund wurde ein Mutationsscreening aller kodierender Exone (Exon 4 – Exon 11) von *GUCY1A3* mittels *high resolution melting analyses* durchgeführt (vgl. 3.7.1). Es wurden 81 MI-Patienten aus den 26 Großfamilien, weitere 483 nicht familiäre MI-Patienten und 562 gesunde Individuen für das Screening ausgewählt (klinische Charakteristika vgl. 8.2 Tabelle 45). Als Positivkontrolle dienten die Individuen der Exom-Sequenzierung mit der *GUCY1A3*-Insertion im Exon 7. In Tabelle 14 ist das Ergebnis des Screenings für die **gesunden Individuen** zusammengefasst. Insgesamt wurden zwei nicht-synonyme Varianten im Exon 7 des Gens identifiziert. Bei diesen handelte es sich um neue Varianten ohne Eintrag in dbSNP132, die mit einer PolyPhen-2 Vorhersage als nicht schädlich für das Protein eingestuft wurden und die an einer nicht konservierten Position des Gens lagen (Tabelle 14).

**Tabelle 14: Ergebnis des Screenings von *GUCY1A3* in gesunden Individuen**

	Annotation NCBI36/hg18 <i>GUCY1A3</i> NM_000856	ID	PolyPhen-2	Konservierung der Mutationsstelle
Exon 7	chr4:156851216: c.A449T, p.N150I	ID 6007703	<i>benign</i>	nein
Exon 7	chr4:156851536: c.G769A, p.V257I	ID 4426701	<i>benign</i>	nein

Bei dem Screening der **HI-Patienten** konnte die Variante der Exom-Sequenzierung in den Individuen der Familie 2001 bestätigt werden (Tabelle 15). Außerdem wurden eine neue synonyme Variante und fünf nicht-synonyme Varianten im *GUCY1A3*-Gen identifiziert. Für die nicht-synonymen Varianten wurde laut PolyPhen-2 ein möglich oder wahrscheinlich schädlicher Einfluss auf die Proteine vorhergesagt. Ein hoher Konservierungsgrad der Mutationsstellen wurde festgestellt. Zwei dieser Varianten, T229M und G537R wurden in Individuen zweier weiterer Großfamilien, Familie 6486 und 8622, gefunden (Tabelle 15).

**Tabelle 15: Ergebnis des Screenings von *GUCY1A3* in HI-Patienten**

	<b>Annotation NCBI36/hg18 <i>GUCY1A3</i> NM_000856</b>	<b>ID</b>	<b>PolyPhen-2</b>	<b>Konservierung der Mutationsstelle</b>
Exon 6	chr4:156848871: c.A351G, p.T117T	7552502	-	nein
Exon 7	chr4:156851453: c.C686T, p.T229M	6486504 (Großfamilie)	<i>possibly damaging</i>	ja
Exon 7	chr4:156851250: c.482_483insT, p.L163Ffs*24	2001511, 2001588, 2001596, 2001552, 2001544 (Großfamilie)	<i>frameshift</i>	ja
Exon 9	chr4:156857899: c.A1711G, p.I571V	9360503	<i>possibly damaging</i>	ja
Exon 9	chr4:156857797: c.G1609A, p.G537R	8622503 (Großfamilie)	<i>probably damaging</i>	ja
Exon 10	chr4:156862682: c.G1759A, p.V587I	8441505	<i>probably damaging</i>	ja
Exon 10	chr4:156862752: c.G1829A, p.C610Y	6510501	<i>probably damaging</i>	ja

Durch das Mutationsscreening wurde deutlich, dass die Wahrscheinlichkeit eine schädliche Variante im Gen *GUCY1A3* zu tragen bei HI-Patienten 5-fach höher ist als bei gesunden Individuen. Außerdem zeigte sich eine Häufung der Varianten in den HI-Großfamilien.

#### 4.1.6. Sequenzierung weiterer Kandidatengene

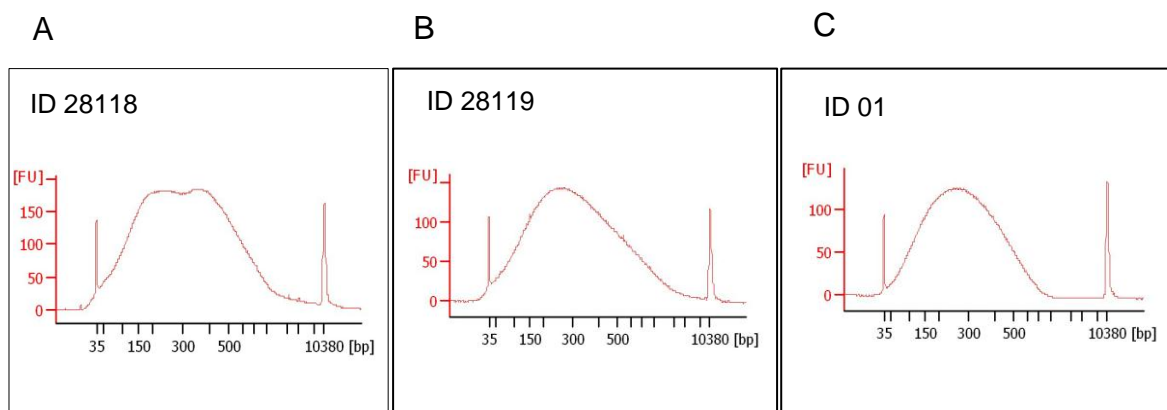
Aufgrund der fehlenden vollständigen Ko-Segregation der Varianten in der Familie 2001, wurde die Hypothese aufgestellt, dass zwei oder mehr Gene über Interaktion der durch die Varianten geschädigten Proteine die Erkrankung KHK/HI in der Familie erklären. Um diese Hypothese zu verfolgen, wurden alle kodierenden Exone von *ETFDH*, *GCLC* und *CCT7* untersucht. Es wurden 13 Exone von *ETFDH*, 16 Exone

---

von *GCLC* und zwölf Exone von *CCT7* in 48 Herzinfarkt-Patienten *forward* und *reverse* sequenziert. Die HI-Patienten wurden aus den HI-Großfamilien ausgewählt (klinische Charakteristika vgl. 8.2 Tabelle 46). Bei den Exonen *ETFDH* Exon 8, *CCT7* Exon 3, *CCT7* Exon 5, *CCT7* Exon 10, *GCLC* Exon 1 und *GCLC* Exon 16 war nur die Sequenzierung in eine Richtung aufgrund von Polynukleotid-*stretches* möglich. Die Auswertung erfolgte mit den einzelsträngigen Sequenzen. Für *GCLC* Exon 9 konnte auch nach Optimierung aller PCR-Bedingungen kein PCR-Produkt amplifiziert werden. In diesem Screening wurden keine neuen, nicht-synonymen Varianten in den Genen identifiziert.

## 4.2. Fragmentierung der DNA für Exom-Sequenzierung DCM- und HCM-Familien

Für die Anreicherung der Exom-Anteile wurde der TruSeq™ Exome Enrichment Kit von Illumina verwendet (vgl. 3.4.2 und 3.4.3). Zur Erstellung der Exom-Bibliotheken wurde die DNA nicht mit dem vorgeschlagenen Covaris-System sondern mit dem Bioruptor durchgeführt. Im ersten Schritt wurde die Fragmentierung optimiert, wobei die gewünschte Fragmentgröße am reproduzierbarsten mit 15 Zyklen erreicht werden konnte (vgl. 3.4.2). In Abbildung 20 sind drei Bioanalyser-Profile fragmentierter DNA für die Exom-Sequenzierung dargestellt. Die Qualität der DNA, ausgedrückt durch den 260/230-Wert von ID 28118 und ID 28119 (DCM-Familie 16725) war bei 1,5 (optimale Qualität 260/230-Wert = 1,8 - 2,0). Die Profile beider DNAs zeigten keine optimale Größenverteilung der fragmentierten DNA, besonders große Fragmente blieben erhalten (Abbildung 20A/B). Bei der Fragmentierung der DNA der ID 01 (DCM-Familie 009-1229) mit einem 260/230-Wert = 1,9, wurde eine optimale Größenverteilung der Fragmente erreicht (Abbildung 20C).

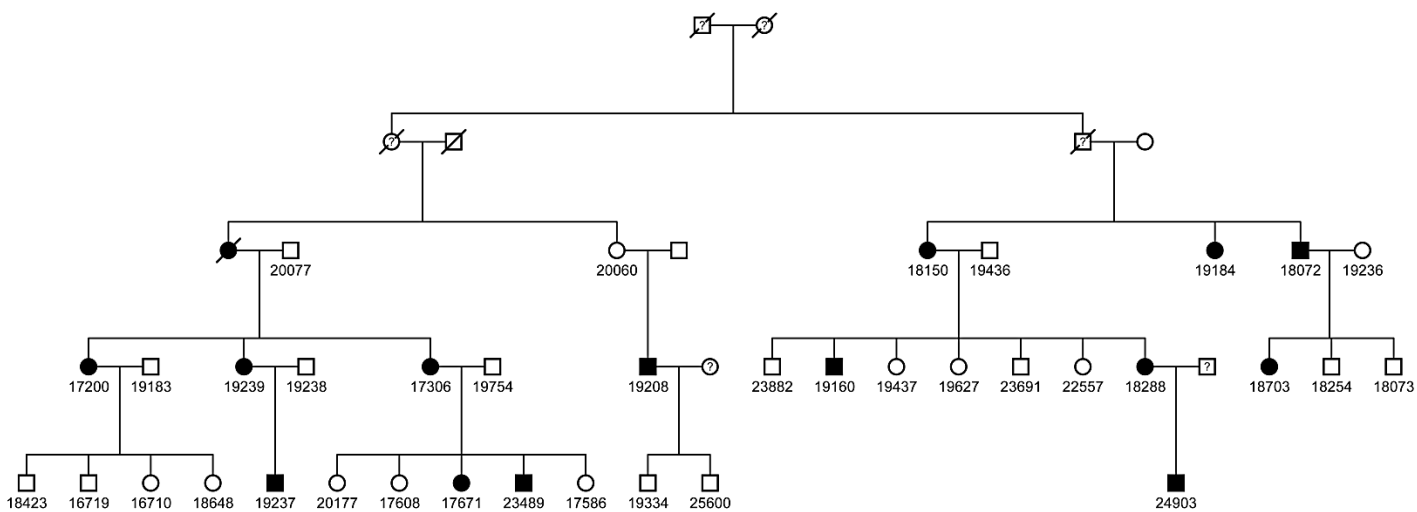


**Abbildung 20: Bioanalyser-Profile fragmentierter DNA. A: ID 18288, B: ID 28118, C: ID 01. Auf der Abszisse ist die Größe der Fragmente in bp aufgetragen, auf der Ordinate die Fluoreszenz in FU.**

Die Sequenzierung erfolgte mit dem HiSeq2000 (vgl. 3.5.1, 3.5.2 und 3.5.4). Die Auswertung der Sequenzen wurde mit der Auswertungsstrategie 2 durchgeführt (vgl. 3.6.3 und 3.6.4).

### 4.3. Exom-Sequenzierung DCM-Familie 9800

Abbildung 21 zeigt den Stammbaum der DCM-Familie 9800. 15 Individuen über drei Generationen erkrankten an DCM. Die klinischen Charakteristika der Familie sind in Tabelle 42 unter 8.1 zusammengefasst.



**Abbildung 21: Stammbaum der DCM-Familie 9800. Quadrat: Mann, Kreis: Frau, ausgefüllte Symbole: betroffen, offene Symbole: nicht betroffen, durchgestrichene Symbole: verstorben, ?: Phänotyp unbekannt.**

Fünf erkrankte Individuen (ID 17200, ID 17671, ID 18288, ID 19208 und ID 19237) wurden für die Exom-Sequenzierung ausgewählt. Die Proben wurden entsprechend dem Standardprotokoll (nicht modifiziert) und mit einem modifizierten Protokoll vorbereitet (vgl. 3.4.3). Die Proben ohne Größenselektion der Fragmente mit einem Agarosegel, d.h. mit dem modifizierten Protokoll, wurden folgend als modifiziert (*modif.*) gekennzeichnet. Bei den nicht modifizierten Proben wurden durchschnittlich 14,1% (12,6% - 16,4%) aller *reads* als Duplikate entfernt (Tabelle 16). Die Anzahl der Duplikate bei den modifizierten Proben war dem gegenüber mit 18,5% (14,7% - 20,5%) erhöht. Dadurch standen weniger *reads* für das *mapping* zur Verfügung (*mapped reads*: 64.938.376 - 86.976.185; *mapped reads* *modif.*: 54.914.863 - 76.248.893), mit Ausnahme von ID 19208 (*mapped reads* ID 19208: 86.501.382; *mapped reads* ID 19208 *modif.*: 88.432.387). Durchschnittlich wurden von den nicht modifizierten Proben 5,8 Gb und von den modifizierten Proben 5,3 Gb der Referenzsequenz zugeordnet. 55% - 63% aller *reads* waren *on target*, wobei

durchschnittlich 96% der Basen zumindest einmal abgebildet wurden. Ungefähr 77% aller Basen erreichten eine *coverage* von  $\geq 20x$  (Tabelle 16).

**Tabelle 16: Information der generierten Exom-Sequenzen für DCM-Familie 9800**

ID	Duplikate in %	reads/ mapped reads	mapped reads in %	mapped sequence in bp	on target in %	coverage $\geq 1x$ in %	coverage $\geq 4x$ in %	coverage $\geq 8x$ in %	coverage $\geq 20x$ in %
17200	13,8	65.522.226/ 64.938.376	99,1	4.870.378.200	63	94	91	88	76
17200 modif.	19,8	60.143.556/ 59.537.882	99,0	4.465.341.150	57	97	93	89	71
17671	14,2	67.673.776/ 67.043.838	99,1	5.028.287.850	60	95	92	89	78
17671 modif.	14,7	55.546.630/ 54.914.863	98,9	4.118.614.725	55	97	93	88	69
18288	16,4	82.775.838/ 81.910.493	99,0	6.143.286.975	59	96	94	91	82
18288 modif.	17,9	72.414.848/ 71.592.930	98,9	5.369.469.750	58	97	94	91	78
19208	12,6	87.317.330/ 86.501.382	99,1	6.487.603.650	56	97	94	91	82
19208 modif.	19,5	89.446.762/ 88.432.387	98,9	6.632.429.025	55	97	95	91	80
19237	13,5	87.846.626/ 86.976.185	99,0	6.523.213.875	56	96	94	91	82
19237 modif.	20,5	77.134.480/ 76.248.893	98,9	5.718.666.975	56	97	94	90	76

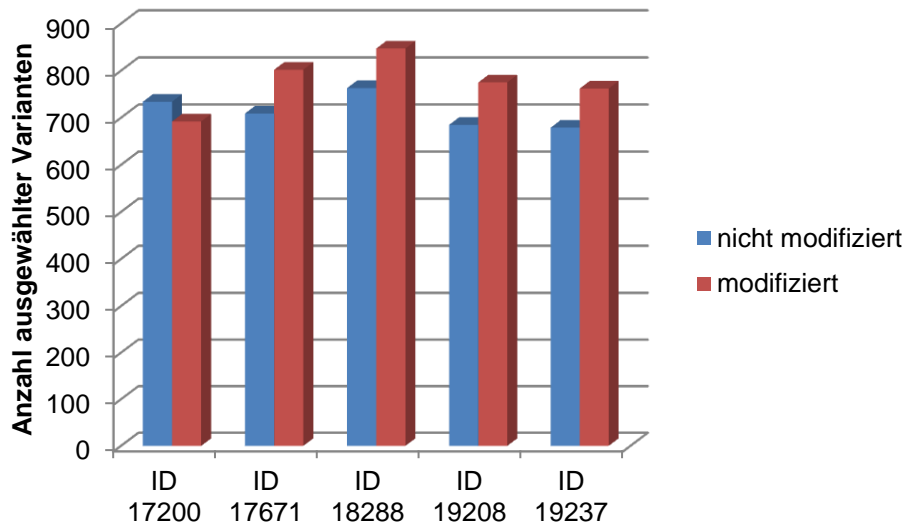
Nach dem *Alignment* der Sequenzen wurden die Varianten detektiert und annotiert. Da ein autosomal-dominanter Erbgang vermutet wurde, wurden von den detektierten Varianten heterozygote nicht-synonyme Varianten, Indels und Spleiß-Varianten ohne rs-Nummer bevorzugt. In Tabelle 17 ist die Anzahl der ausgewählten Varianten für die Individuen der DCM-Familie 9800 dargestellt. Eine Gesamtübersicht der detektierten Varianten findet sich in Tabelle 52 unter 8.6.

**Tabelle 17: Anzahl ausgewählter Varianten der DCM-Familie 9800**

ID	Anzahl heterozygoter nicht-synonymer Varianten, Indels und Spleiß-Varianten ohne rs-Nummer	
	nicht modifiziert	modifiziert
17200	734	692
17671	709	802
18288	763	847
19208	685	775
19237	679	762

Pro Individuum wurde bei modifizierter und nicht modifizierter Probe eine vergleichbare Anzahl an ausgewählten Varianten detektiert. Für ID 17200 modif.

wurden tendenziell weniger Varianten ( $n=692$ ) detektiert als bei der nicht modifizierten Probe ( $n=734$ ). Bei den anderen Proben war die Tendenz in die andere Richtung, d.h. es wurden mehr Varianten bei den modifizierten Proben ( $n=775$  bis  $n=847$ ) als bei den nicht modifizierten Proben ( $n=679$  bis  $n=763$ ) detektiert (Tabelle 17, Abbildung 22). In Abbildung 22 ist die Anzahl der ausgewählten Varianten für alle IDs aufgetragen.



**Abbildung 22: Anzahl ausgewählter Varianten der DCM-Familie 9800. Es wurde die Anzahl ausgewählter Varianten der nicht modifizierten und modifizierten Proben pro ID aufgetragen. Der Vergleich zeigte, dass beide Protokolle zu einer vergleichbaren Anzahl an detektierten Varianten führten.**

Nach der Identifizierung der ausgewählten Varianten, wurden diese zwischen den fünf sequenzierten Familienmitgliedern verglichen. Bei dem Vergleich wurde zwischen modifizierter und nicht modifizierter Probe unterschieden. Für die von den fünf Individuen geteilten Varianten wurde ein zusätzlicher Filter geschaltet (vgl. 3.6.5). Paraloge Sequenzen im Exom, die durch Genduplikation entstanden sind, führten zu Problemen bei der Detektion von Varianten. Ursprünglich verschiedene Regionen wurden beim *mapping* einer gleichen Position zugeordnet, wodurch falsch positive Varianten erzeugt wurden. Alle Genbereiche, die eine Variante beinhalten, wurden deshalb auf Paralogie untersucht und danach gefiltert. In Tabelle 18 sind die Genregionen der DCM-Familie 9800 (modifiziert/nicht modifiziert) gezeigt, die als paralog und damit als falsch positiv eingestuft wurden. Vertreter der Gen-Familie

*ANKRD* (ankyrin repeat domain family) und *FAM* (family with sequence similarity) waren mehrheitlich vertreten (Tabelle 18).

**Tabelle 18: Paraloge Genregionen der DCM-Familie 9800**

Position	Gen	Paraloge Gene
chr1:17030477	<i>ESPNP</i>	<i>ESPN</i>
chr1:201179613	<i>IGFN1</i>	<i>IGNF1</i>
chr2:132905792	<i>ANKRD30BL</i>	<i>ANKRD30BP2</i>
chr3:195507062	<i>MUC4</i>	<i>MUC4</i>
chr9:67272289	<i>AQP7P1</i>	<i>AQP7P1</i>
chr9:69663350	<i>LOC100133920</i>	<i>LOC286297</i>
chr13:19409574	<i>ANKRD20A9P</i>	<i>ANKRD20A5P</i>
chr13:19412955	<i>ANKRD20A9P</i>	<i>ANKRD20A5P</i>
chr16:22545897	<i>LOC100132247</i>	<i>LOC613037</i>
chr20:25753403	<i>FAM182B</i>	<i>FAM182A</i>
chr20:26064076	<i>FAM182A</i>	<i>FAM182B</i>
chr20:26064328	<i>FAM182A</i>	<i>FAM182B</i>
chr20:26064330	<i>FAM182A</i>	<i>FAM182B</i>
chr20:26064408	<i>FAM182A</i>	<i>FAM182B</i>
chr21:9908223	<i>TEKT4P2</i>	<i>LOC100132288, FLJ00219</i>

Varianten in nicht paralogen Genregionen wurden weiter betrachtet. In Tabelle 19 sind die Varianten aufgelistet, die aus dem Vergleich der nicht modifizierten Proben resultierten. Insgesamt wurden drei Varianten von den fünf sequenzierten Familienmitgliedern geteilt. Eine Variante war in dem Gen *LOC645752* lokalisiert. Dabei handelte es sich um eine nicht-kodierende RNA (*noncoding RNA*) mit einer Variante an einer nicht konservierten Position. In Tabelle 19 wurde das Fehlen der Konservierung mit „nein“ gekennzeichnet. Zwei der drei Varianten waren im Gen *SEPHS1P* lokalisiert, das ebenso für eine nicht-kodierende RNA kodiert. Die Konservierung der Mutationsstellen war nicht gegeben (Tabelle 19).

**Tabelle 19: Geteilte Varianten der Exome von DCM-Familie 9800 (nicht modifiziert)**

Gen	Annotation GRCh37/hg19	PolyPhen-2	Konservierung der Mutationsstelle
1. <i>LOC645752</i>	NR_027024: Position chr15:78211344: Exon 11: c.G1187A	<i>noncoding RNA</i>	nein
2. <i>SEPHS1P</i>	NR_002789: Position chr7:64312104: Exon 1: c.2074_0del	<i>noncoding RNA</i>	nein
3. <i>SEPHS1P</i>	NR_002789: Position chr7:64313689: Exon 1: c.A490C	<i>noncoding RNA</i>	nein



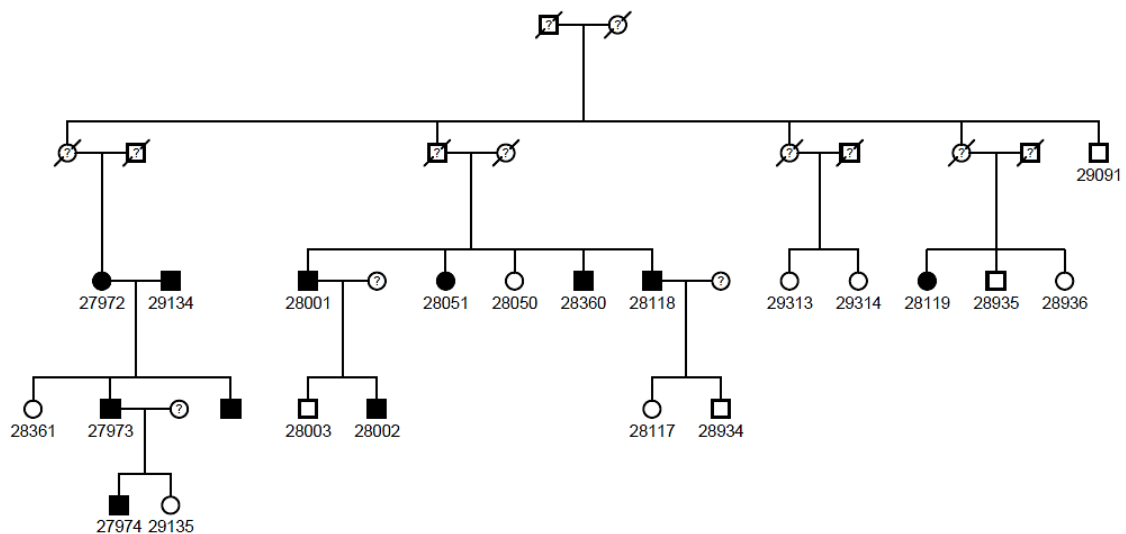
Aus dem Vergleich der Varianten der modifizierten Proben resultierten sechs Varianten (Tabelle 20). Die drei Varianten aus dem Vergleich der nicht modifizierten Proben wurden ebenfalls identifiziert. Zusätzlich wurden zwei Varianten in nicht-kodierende RNAs, *AK2* und *LOC728725* und eine Variante im Gen *PASD1* identifiziert. Die Variante im Gen *PASD1* wurde mit PolyPhen-2 als möglich schädlich für das entstehende Protein eingestuft. Außerdem wurde die Konservierung der Mutationsstelle betrachtet. Diese war entweder gegeben (ja), nicht gegeben (nein) oder durch fehlende Sequenzen von anderen Organismen nicht bekannt (nicht bekannt). Die Mutationsstelle für die Variante im Gen *PASD1* war nicht konserviert (Tabelle 20).

**Tabelle 20: Geteilte Varianten der Exome von DCM-Familie 9800 (modifiziert)**

Gen	Annotation GRCh37/hg19	PolyPhen-2	Konservierung der Mutationsstelle
1. <i>LOC645752</i>	NR_027024: Position chr15:78211344: Exon 11: c.G1187A	<i>noncoding RNA</i>	nein
2. <i>SEPHS1P</i>	NR_002789: Position chr7:64312104: Exon 1: c.2074_0del	<i>noncoding RNA</i>	nein
3. <i>SEPHS1P</i>	NR_002789: Position chr7:64313689: Exon 1: c.A490C	<i>noncoding RNA</i>	nein
4. <i>AK2</i>	NR_037591: Position chr1:33476226: Exon 8: c.1103_0del	<i>noncoding RNA</i>	ja
5. <i>LOC728758</i>	NR_003571: Position chr15:44168026: Exon 1: c.G2344A	<i>noncoding RNA</i>	nicht bekannt
6. <i>PASD1</i>	NM_173493: Position chrX:150791434: Exon 7: c.G444T, p.L148F	<i>possibly damaging</i>	nein

#### 4.4. Exom-Sequenzierung DCM-Familie 16725

Abbildung 23 zeigt den Stammbaum der DCM-Familie 16725. Zehn Familienmitglieder waren an DCM erkrankt. Zudem wies der Ehemann (ID 29134) von ID 27972 eine DCM-Erkrankung auf. Die klinischen Charakteristika der Familie sind in Tabelle 40 unter 8.1 aufgeführt.



**Abbildung 23: Stammbaum der DCM-Familie 16725. Quadrat: Mann, Kreis: Frau, ausgefüllte Symbole: betroffen, offene Symbole: nicht betroffen, durchgestrichene Symbole: verstorben, ?: Phänotyp unbekannt.**

Für die Exom-Sequenzierung wurden vier betroffene Familienmitglieder (ID 27974, ID 28002, ID 28118, ID 28119) und der betroffene Nicht-Familienangehörige (ID 29134) ausgewählt. Pro Individuum wurden beide Protokolle, modifiziert und nicht modifiziert angewendet. In Tabelle 21 ist die Information der generierten Exom-Sequenzen dargestellt. Bei der nicht modifizierten Probe der ID 27974 zeigte sich mit 28,3% eine sehr hohe Anzahl an Duplikaten gegenüber den weiteren nicht modifizierten Proben mit 11,8% - 15,7%. Die Anzahl der Duplikate war mit Ausnahme der ID 27974 bei den modifizierten Proben im Vergleich zu den nicht modifizierten Proben mit 14,1% - 18,0% erhöht. Durchschnittlich wurden für alle Proben 81% der *reads* (84,7% - 98,1%) zugeordnet, wobei pro Individuum eine vergleichbare Prozentzahl beobachtet wurde. In früheren Arbeiten wurde gezeigt, dass pro Exom 6 Gb zugeordneter Sequenz zu erwarten sind (vgl. 3.6.1). Von den dargestellten Exom-Sequenzen wurde dies nur annähernd bei der Probe mit der ID 29134

beobachtet. Vor allem für ID 27974<sub>modif.</sub> mit ~ 532 Kb, für ID 28118<sub>modif.</sub> mit ~ 371 Kb und für ID 28119<sub>modif.</sub> mit ~ 521 Kb zugeordneter Sequenz wurde dieser Bereich stark unterschritten. Die geringe Menge an generierte Sequenzen spiegelte sich in der *coverage* wider. Bei den genannten Proben wurden 64% - 80% aller Basen einmal abgebildet. Eine *coverage* von  $\geq 8x$  wurde nur von 2% - 5% aller Basen erreicht. Eine wünschenswerte *coverage* von  $\geq 20x$  konnte nicht abgebildet werden (Tabelle 21).

**Tabelle 21: Information der generierten Exom-Sequenzen für DCM-Familie 16725**

ID	Duplikate in %	reads/ mapped reads	mapped reads in %	mapped sequence in bp	on target in %	coverage $\geq 1x$ in %	coverage $\geq 4x$ in %	coverage $\geq 8x$ in %	coverage $\geq 20x$ in %
27974	28,3	32.491.714/ 28.086.997	86,4	2.106.524.775	54	93	77	51	8
27974 modif.	19,2	8.037.210/ 7.098.254	88,3	532.369.050	53	80	29	5	0
28002	11,8	47.990.662/ 47.067.780	98,1	3.530.083.500	56	96	92	86	62
28002 modif.	14,1	22.932.900/ 22.448.050	97,9	1.683.603.750	55	95	85	67	17
28118	15,7	16.933.940/ 14.341.229	84,7	1.075.592.175	50	86	51	21	1
28118 modif.	18,0	5.771.714/ 4.950.416	85,8	371.281.200	50	64	14	2	0
28119	12,9	24.970.618/ 21.547.520	86,3	1.616.064.000	51	92	73	44	6
28119 modif.	14,3	7.989.390/ 6.950.472	87,0	521.285.400	51	78	27	5	0
29134	14,2	55.359.378/ 53.926.979	97,4	4.044.523.425	58	96	92	88	68
29134 modif.	16,3	30.220.434/ 29.373.717	97,2	2.203.028.775	55	96	88	76	29

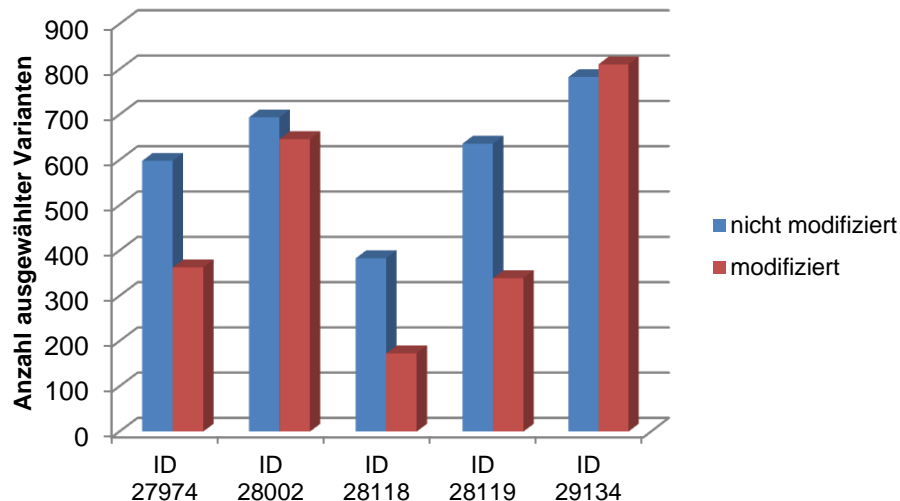
Varianten der Exom-Sequenzen wurden detektiert und annotiert. Ebenso wie bei der DCM-Familie 9800 wurden heterozygote nicht-synonyme Varianten, Indels und Spleiß-Varianten ohne rs-Nummer bevorzugt. In Tabelle 22 ist die Anzahl der ausgewählten Varianten pro Individuum dargestellt (vgl. 8.6 Tabelle 53). Durch vorausgegangene Publikationen war bekannt, dass pro Exom im Durchschnitt 500-700 ausgewählte Varianten zu erwarten sind (vgl. 3.6.1). Die Anzahl der ausgewählten Varianten korreliert mit der Menge an generierter Sequenz-Information. Für die nicht modifizierten und modifizierten Proben der ID 28002 (n=694 bzw. n=646) und der ID 29134 (n=783 bzw. n=811) lag die Anzahl im erwarteten Bereich, ebenso für die nicht modifizierten Proben der ID 27974 (n=598) und der ID 28119 (n=636). Die weiteren Proben dieser Familie mit mangelnder Sequenz-Information zeigten eine geringere Anzahl mit n=173 bis n=383 an ausgewählten Varianten (Tabelle 22).

**Tabelle 22: Anzahl ausgewählter Varianten der DCM-Familie 16725**

ID	Anzahl heterozygoter nicht-synonymer Varianten, Indels und Spleiß-Varianten ohne rs-Nummer	
	nicht modifiziert	modifiziert
27974	598	363
28002	694	646
28118	383	173
28119	636	339
29134	783	811

#### 4.4.1. Identifizierte Varianten

Für die weitere Auswertung wurden die Varianten zwischen den Familienmitgliedern verglichen. Da die Person mit der ID 29134 kein Familienmitglied war, wurden ausschließlich Varianten ausgesucht, die in ID 29134 nicht vorkamen, aber die von den Familienmitgliedern geteilt wurden. Wie Tabelle 22 zeigt und in Abbildung 24 nochmals veranschaulicht wird, war die Anzahl der ausgewählten Varianten bei den modifizierten Proben nur für ID 28002 und ID 29134 im erwarteten Bereich. Deshalb wurde die Auswertung für die modifizierten Proben nicht durchgeführt. Bei der Auswertung der nicht modifizierten Proben wurde die Probe der ID 28118 ausgeschlossen. Abbildung 24 zeigt die verhältnismäßig kleine Anzahl an ausgewählten Varianten für ID 28118, die deshalb nicht betrachtet wurden. Zusammenfassend wurden für die Auswertung der nicht modifizierten Proben Varianten ausgewählt, die in ID 27974, ID 28002 und ID 28119 und nicht in ID 29134 gefunden wurden (Tabelle 22/Abbildung 24).



**Abbildung 24: Anzahl ausgewählter Varianten der DCM-Familie 16725. Es wurde die Anzahl ausgewählter Varianten der nicht modifizierten und modifizierten Proben pro ID aufgetragen. Der Vergleich zeigte große Unterschiede zwischen den IDs.**

Nachdem die Auswertung der Varianten erfolgt war, wurden die Genregionen der Varianten auf Paralogie untersucht. Es wurden unter anderem zwei Varianten im Gen *MUC4* und im Gen *FAM182A* als paralog eingestuft und von der weiteren Betrachtung ausgeschlossen (Tabelle 23).

**Tabelle 23: Paraloge Genregionen der DCM-Familie 16725**

Position	Gen	Paraloge Gene
chr1:14464	<i>WASH7P</i>	<i>WASH3P</i>
chr3:195507062	<i>MUC4</i>	<i>MUC4</i>
chr3:195507683	<i>MUC4</i>	<i>MUC4</i>
chr19:40392802	<i>FCGBP</i>	<i>FCBBP</i>
chr20:26064408	<i>FAM182A</i>	<i>FAM182B</i>
chr20:26067039	<i>FAM182A</i>	<i>FAM182B</i>

Insgesamt wurden acht geteilte Varianten als nicht paralog eingestuft (Tabelle 24). Drei dieser Varianten befanden sich in Genen (*LOC650623*, *FLJ43681*), die für nicht-kodierende RNAs kodierten. Eine Variante wurde im Gen *MUC4* identifiziert, die nicht weiter beschrieben werden konnte, da mit PolyPhen-2 keine Vorhersage getroffen werden konnte (*unknown*) und Sequenzen anderer Organismen fehlten, um die Konservierung zu betrachten. Des Weiteren wurden zwei Varianten in den Genen *CAP1* und *ATG16L2* von den drei betroffenen Familienmitgliedern geteilt, die laut PolyPhen-2 nicht schädlich für die entstehenden Proteine waren. Eine Deletion von drei Basenpaaren, die somit nicht zu einer Veränderung des Leserahmens führte

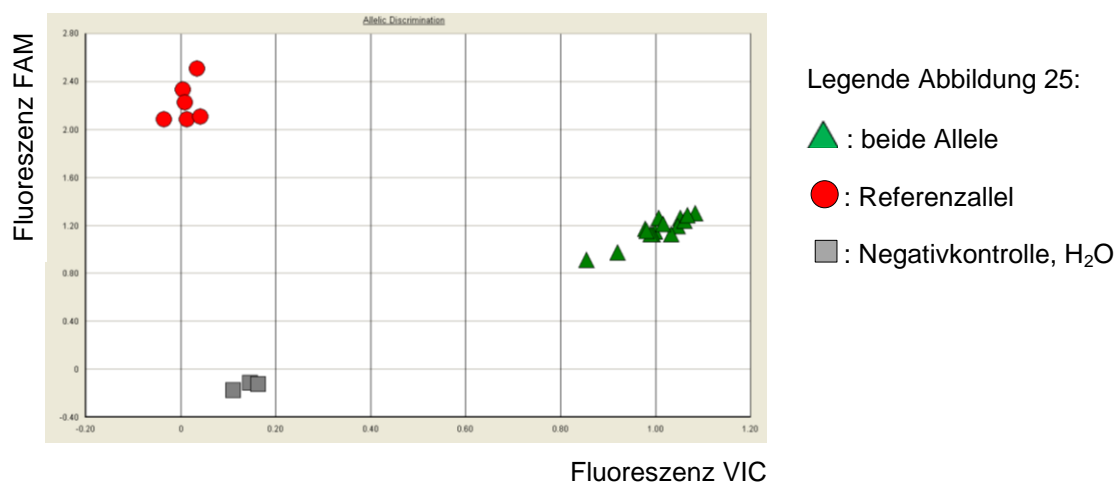
(*nonframeshift*), wurde im Gen *KCNJ12* detektiert. Außerdem wurde im Titin-Gen (*TTN*) eine Variante identifiziert durch die ein Stoppkodon in der Sequenz eingeführt wurde (*stopgain*). Die Mutationsstellen der zwei letztgenannten Varianten waren konserviert (Tabelle 24). *KCNJ12* und *TTN* stellen plausible Kandidatengene dar (siehe 5.4).

**Tabelle 24: Geteilte Varianten der Exome von DCM-Familie 16725, die nicht in ID 29134 vorkamen**

Gen	Annotation GRCh37/hg19	PolyPhen-2	Konservierung der Mutationsstelle
1. <i>LOC650623</i>	NR_027512: Position chr10:81446059: Exon 1: c.T3329C	<i>noncoding RNA</i>	nicht bekannt
2. <i>LOC650623</i>	NR_027512: Position chr10:81446886: Exon 1: c.T4156C	<i>noncoding RNA</i>	nicht bekannt
3. <i>FLJ43681</i>	NR_029406: Position chr17:81177189: Exon 3: c.G901T	<i>noncoding RNA</i>	nicht bekannt
4. <i>MUC4</i>	NM_018406: Position chr3:195506473: Exon 2: c.A11978C; p.H3993P	<i>unknown</i>	nicht bekannt
5. <i>CAP1</i>	NM_006367: Position chr1:40535490: Exon 9: c.C937G; p.R313G	<i>benign</i>	nein
6. <i>ATG16L2</i>	NM_033388: Position chr11:72528866: Exon 3: c.A284G; p.E95G	<i>benign</i>	nein
7. <i>KCNJ12</i>	NM_021012: Position chr17:21319654: Exon 3: c.1000_1002del; p.334_334del	<i>nonframeshift</i>	ja
8. <i>TTN</i>	NM_003319: Position chr2:179411847: Exon 168: c.A67210T; p.K22404X	<i>stopgain</i>	ja

#### 4.4.2. Ko-Segregation der *TTN*-Variante

Die Ko-Segregation der Variante im Titin-Gen wurde durch Genotypisierung untersucht (vgl. 3.7.2). Alle Familienmitglieder und ID 29134 wurden genotypisiert. Das Ergebnis ist in Abbildung 25 dargestellt. FAM-Fluoreszenz entsteht durch das Vorhandensein des Referenzallels. Fluoreszenz sowohl vom Farbstoff FAM als auch VIC entsteht durch das Vorhandensein von Referenzallel/mutiertes Allel. Die Genotypisierung zeigte, dass sechs Individuen Träger des Referenzallels und 15 Individuen heterozygot für die Titin-Variante waren (Abbildung 25).



**Abbildung 25: Genotypisierung der *TTN*-Variante in der DCM-Familie 16725. Auf der Abszisse ist die Fluoreszenz von VIC angegeben, auf der Ordinate die Fluoreszenz von FAM. Als Negativkontrolle wurde H<sub>2</sub>O verwendet.**

In Tabelle 25 ist das Ergebnis der Genotypisierung für alle IDs zusammengefasst. Es konnte bestätigt werden, dass ID 29134 als Nicht-Familienmitglied kein Träger der *TTN*-Variante ist. ID 28118 wurde von der Auswertung der Exom-Sequenzen wegen mangelnder Sequenz-Information ausgeschlossen. Durch die Genotypisierung wurde in dieser ID sowie für vier weitere DCM-Betroffene (ID 27972, ID 28001, ID 28360, ID 27973) die *TTN*-Variante identifiziert. Sieben der elf Nicht-Betroffenen der Familie waren ebenso Träger der Variante (Tabelle 25).

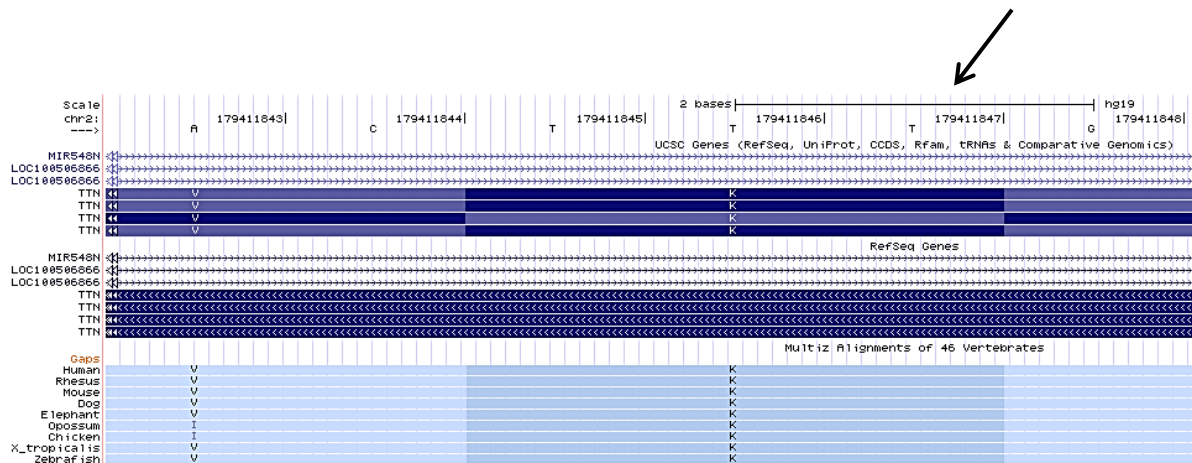
**Tabelle 25: Ko-Segregation der *TTN*-Variante in der DCM-Familie 16725**

ID	Phänotyp *	<i>TTN</i> -Variante
29091	Nicht betroffen, 79 Jahre	negativ
27972	DCM, 69 Jahre	positiv
29134/kein Familienmitglied	DCM, 74 Jahre	negativ
28001	DCM, 49 Jahre	positiv
28051	DCM, 68 Jahre	negativ
28050	Nicht betroffen, 66 Jahre	positiv
28360	DCM, 57 Jahre	positiv
28118	DCM, 48 Jahre	positiv
29313	Nicht betroffen, 53 Jahre	positiv
29314	Nicht betroffen, 57 Jahre	positiv
28119	DCM, 69 Jahre	positiv
28935	Nicht betroffen, 71 Jahre	negativ
28936	Nicht betroffen, 65 Jahre	positiv
28361	Nicht betroffen, 36 Jahre	positiv
27973	DCM, 38 Jahre	positiv
28003	Nicht betroffen, 19 Jahre	positiv
28002	DCM, 25 Jahre	positiv
28117	Nicht betroffen, 34 Jahre	positiv
28934	Nicht betroffen, 31 Jahre	negativ
27974	DCM, 21 Jahre	positiv
29135	Nicht betroffen, 17 Jahre	negativ
Familienmitglieder	Betroffene mit Mutation/ Gesamtzahl Betroffene	8/9
Familienmitglieder	Nicht-Betroffene mit Mutation/ Gesamtzahl Nicht-Betroffene	7/11

\* Bei Betroffenen Diagnosealter, bei Unbetroffenen Einschlussalter



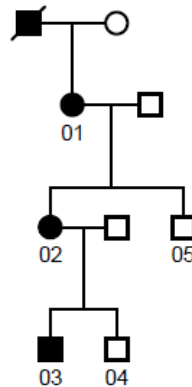
Die Mutationsstelle im Titin-Gen, die zur Einführung des Stoppkodons führte, zeigte einen hohen Konservierungsgrad. Für alle gezeigten Organismen (Wirbeltiere) kodierten die drei Basenpaare in der Sequenz für die Aminosäure Lysin (Abbildung 26).



**Abbildung 26: Konservierung der Mutationsstelle der TTN-Variante. Der Pfeil markiert die Position der Mutationsstelle im Gen TTN. Die Konservierung ist für folgende Organismen gezeigt: Mensch (*human*), Rhesusaffe (*rhesus*), Maus (*mouse*), Hund (*dog*), Elefant (*elephant*), Opossum (*opossum*), Huhn (*chicken*), Frosch (*Xenopus tropicalis*) und Fisch (*zebrafish*) (<http://genome.ucsc.edu/>).**

#### 4.5. Exom-Sequenzierung DCM-Familie 009-1229

Bei der DCM-Familie 009-1229 handelte es sich um eine kleinere Familie mit vier an DCM erkrankten Individuen (Abbildung 27). Die klinischen Charakteristika sind in Tabelle 41 unter 8.1 aufgelistet.



**Abbildung 27: Stammbaum der DCM-Familie 009-1229. Quadrat: Mann, Kreis: Frau, ausgefüllte Symbole: betroffen, offene Symbole: nicht betroffen, durchgestrichene Symbole: verstorben.**

Von zwei betroffenen Familienmitgliedern mit ID 01 und ID 03 wurden die Proben mit beiden Protokollen vorbereitet und das Exom sequenziert. Bei drei Proben wurden 11% - 15% aller *reads* als Duplikate entfernt (Tabelle 26). Nur bei der Probe der ID 03<sub>modif.</sub> war die Anzahl der Duplikate mit 20% erhöht. Für alle Proben wurde eine große Anzahl an *reads* (60.748.362 – 69.162.862) generiert, von denen >99% zugeordnet werden konnten. Durchschnittlich waren 59% der 4,5 Gb – 5,1 Gb der zugeordneten Sequenz *on target*. Es wurde bei allen Proben eine hohe *coverage* beobachtet. 74% - 79% aller Basen erreichten eine *coverage* von  $\geq 20x$  (Tabelle 26).

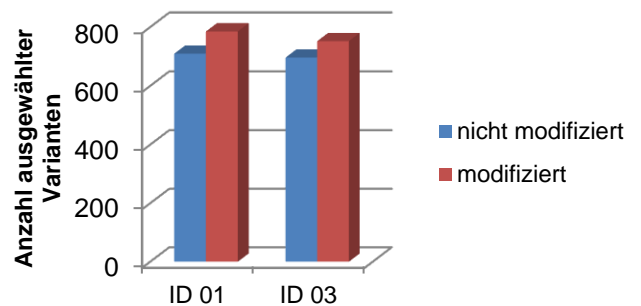
**Tabelle 26: Information der generierten Exom-Sequenzen für DCM-Familie 009-1229**

ID	Duplikate in %	<i>reads/mapped reads</i>	<i>mapped reads</i> in %	<i>mapped sequence</i> in bp	<i>on target</i> in %	<i>coverage</i> $\geq 1x$ in %	<i>coverage</i> $\geq 4x$ in %	<i>coverage</i> $\geq 8x$ in %	<i>coverage</i> $\geq 20x$ in %
01	11,9	60.748.362/ 60.256.456	99,2	4.519.234.200	61	96	92	89	76
01 modif.	14,8	64.294.816/ 63.717.016	99,1	4.778.776.200	57	97	93	90	75
03	11,4	69.162.862/ 68.606.086	99,2	5.145.456.450	59	96	93	90	79
03 modif.	20,3	65.872.378/ 65.318.543	99,2	4.898.890.725	58	97	94	90	74

Pro Individuum und Protokoll konnte eine vergleichbare Anzahl an Varianten nachgewiesen werden (vgl. 8.6 Tabelle 54). Bei den modifizierten Proben wurden mehr ausgewählte Varianten (heterozygote nicht-synonyme Varianten, Indels und Spleiß-Varianten ohne rs-Nummer) detektiert, als bei den nicht modifizierten Proben. Mit  $n=696$  bis  $n=785$  lag die Anzahl der ausgewählten Varianten pro Probe im erwarteten Bereich (Tabelle 27, Abbildung 28).

**Tabelle 27: Anzahl ausgewählter Varianten der DCM-Familie 009-1229**

ID	Anzahl heterozygoter nicht-synonymer Varianten, Indels und Spleiß-Varianten ohne rs-Nummer	
	nicht modifiziert	modifiziert
01	709	785
03	696	752



**Abbildung 28: Anzahl ausgewählter Varianten der DCM-Familie 009-1229. Es wurde die Anzahl ausgewählter Varianten der nicht modifizierten und modifizierten Proben pro ID aufgetragen. Der Vergleich zeigte, dass beide Protokolle zu einer vergleichbaren Anzahl an detektierten Varianten führten.**

#### 4.5.1. Identifizierte Varianten der modifizierten und nicht modifizierten Proben

Die Auswertung der ausgewählten Varianten wurde für beide Protokolle getrennt durchgeführt. Es wurden die Varianten zwischen ID 01 und ID 03 verglichen. 53 Varianten in paralogen Genregionen wurden identifiziert und gefiltert. Neben Vertretern der Gen-Familie *ANKRD* und *FAM* wurden auch Vertreter der Gen-Familie *GOLGA* (*golgin subfamily A*) mehrheitlich von der weiteren Betrachtung ausgeschlossen (Tabelle 28).

Tabelle 28: Paralogie Genregionen der DCM-Familie 009-1229

Position	Gen	Paraloge Gene
chr1:144814702	<i>NBPF9</i>	<i>NBPF15</i>
chr1:144814809	<i>NBPF9</i>	<i>NBPF11</i>
chr1:145302745	<i>NBPF10</i>	<i>NBPF11</i>
chr1:149287051	<i>LOC388692</i>	<i>LOC728855, FAM91A2</i>
chr1:149287492	<i>LOC388692</i>	<i>LOC728855</i>
chr1:248801954	<i>OR2T35</i>	<i>OR2T2</i>
chr2:114369579	<i>RPL23AP7</i>	<i>MGC2752</i>
chr2:87276729	<i>LOC285074</i>	<i>ANAPC1</i>
chr2:98128194	<i>ANKRD36B</i>	<i>ANKRD36</i>
chr2:98128196	<i>ANKRD36B</i>	<i>ANKRD36</i>
chr3:195507683	<i>MUC4</i>	<i>MUC4</i>
chr3:195510217	<i>MUC4</i>	<i>MUC4</i>
chr3:195510310	<i>MUC4</i>	<i>MUC4</i>
chr3:197349143	<i>LOC220729</i>	<i>SDHAP1, SDHAP2</i>
chr4:190878563	<i>FRG1</i>	<i>FRG1B</i>
chr7:102212939	<i>POLR2J3</i>	<i>POLR2J2</i>
chr7:143956432	<i>LOC728377, OR2A7</i>	<i>OR2A4</i>
chr7:144062632	<i>ARHGEF5</i>	<i>LOC728377</i>
chr7:151932945	<i>MLL3</i>	<i>AK310441</i>
chr7:151945072	<i>MLL3</i>	<i>BAGE2-5</i>
chr7:6791040	<i>PMS2CL</i>	<i>PMS2</i>
chr7:74312339	<i>PMS2P5</i>	<i>PMS2L2, PMS2L14</i>
chr7:76682191	<i>LOC100132832</i>	<i>SPDYE1</i>
chr8:11995540	<i>USP17L2</i>	<i>FAM66B</i>
chr8:12040425	<i>FAM86B1</i>	<i>FAM86HP</i>
chr9:141070687	<i>TUBBP5</i>	<i>TUBB4Q</i>
chr9:40702760	<i>FAM75A3</i>	<i>FAM75A6</i>
chr9:69391174	<i>ANKRD20A4</i>	<i>ANKRD20A3, ANKRD20A1</i>
chr9:69664022	<i>LOC100133920</i>	<i>LOC286297</i>
chr9:88431019	<i>LOC389765</i>	<i>KIF27</i>
chr10:46321904	<i>AGAP4</i>	<i>AGAP5</i>
chr11:1264742	<i>MUC5B</i>	<i>MUC5B</i>
chr12:11546194	<i>PRB2</i>	<i>PRB2</i>
chr12:8353318	<i>FAM66C</i>	<i>FAM66B</i>
chr12:90902	<i>LOC100288778</i>	<i>WASH3P</i>
chr13:19414201	<i>ANKRD20A9P</i>	<i>ANKRD20A5P</i>
chr15:20743586	<i>GOLGA6L6</i>	<i>GOLGA6L1</i>
chr15:20770564	<i>GOLGA8C</i>	<i>GOLGA8C</i>
chr15:20771609	<i>GOLGA8C</i>	<i>GOLGA8G, GOLGA8F</i>
chr15:20776622	<i>GOLGA8C</i>	<i>GOLGA8DP</i>
chr15:20777925	<i>GOLGA8C</i>	<i>GOLGA8G, GOLGA8F</i>
chr15:22706758	<i>GOLGA8DP</i>	<i>GOLGA8C</i>
chr16:21416094	<i>NPIPL3</i>	<i>LOC613037</i>
chr16:70011737	<i>PDXDC2P</i>	<i>NPIPL1</i>
chr19:22574988	<i>ZNF98</i>	<i>ZNF492</i>
chr19:22586310	<i>ZNF98</i>	<i>ZNF492</i>
chr20:25753377	<i>FAM182B</i>	<i>FAM182A</i>
chr20:26054459	<i>FAM182A</i>	<i>FAM182B</i>
chr20:26064428	<i>FAM182A</i>	<i>FAM182B</i>
chr21:9908154	<i>TEKT4P2</i>	<i>MAFIP</i>
chr22:18766357	<i>GGT3P</i>	<i>GGT1</i>
chr22:21638248	<i>POM121L8P</i>	<i>POM121L7</i>
chr22:21646811	<i>POM121L8P</i>	<i>POM121L10P</i>

In Tabelle 29 sind die Varianten aufgelistet, die sowohl bei den modifizierten Proben als auch bei den nicht modifizierten Proben identifiziert wurden. Insgesamt wurden 104 Varianten von ID 01/ID 01<sub>modif.</sub> und ID 03/ID 03<sub>modif.</sub> geteilt. Von diesen Varianten lagen 76 im kodierenden Bereich mit drei Indels und 73 SNVs. Mit PolyPhen-2 wurden 28 SNVs als möglich oder wahrscheinlich schädlich für die entstehenden Proteine klassifiziert. Weiterhin wurden 28 Varianten im nicht-kodierenden Bereich detektiert. Für alle Varianten wurde die Konservierung der Mutationsstelle betrachtet (Tabelle 29). Diese war entweder gegeben (ja), nicht gegeben (nein), durch fehlende Sequenzen von anderen Organismen nicht bekannt (nicht bekannt) oder nur in manchen der gelisteten Organismen vorhanden (nicht vollständig).

**Tabelle 29: Geteilte Varianten der Exome von DCM-Familie 009-1229 (modifiziert/nicht modifiziert)**

Gen	Annotation GRCh37/hg19	PolyPhen-2	Konservierung der Mutationsstelle
1. <i>GJB4</i>	NM_153212: Position chr1:35227010: Exon 2: c.155_158del, p.52_53del	<i>frameshift</i>	nicht vollständig
2. <i>DEFB126</i>	NM_030931: Position chr20:126160: Exon 2: c.163_166del, p.55_56del	<i>frameshift</i>	nein
3. <i>DLEC1</i>	NM_007337: Position chr3:38164073: Exon 36: c.5314_5316del, p.1772_1772del	<i>nonframeshift</i>	nicht vollständig
4. <i>ACTA1</i>	NM_001100: Position chr1:229568161: Exon 4: c.G472A, p.G158S	<i>possibly damaging</i>	ja
5. <i>NUBP2</i>	NM_012225: Position chr16:1837991: Exon 5: c.C539T, p.T180M	<i>possibly damaging</i>	ja
6. <i>CAPN12</i>	NM_144691: Position chr19:39221828: Exon 19: c.A1993C, p.T665P	<i>possibly damaging</i>	ja
7. <i>TPX2</i>	NM_012112: Position chr20:30345332: Exon 3: c.A53G, p.N18S	<i>possibly damaging</i>	ja
8. <i>FAM83F</i>	NM_138435: Position chr22:40415250: Exon 2: c.G568A, p.V190I	<i>possibly damaging</i>	ja
9. <i>TUBGCP6</i>	NM_020461: Position chr22:50667930: Exon 4: c.C1193T, p.S398L	<i>possibly damaging</i>	ja
10. <i>HMGCLL1</i>	NM_019036: Position chr6:55304355: Exon 9: c.G888A, p.M296I	<i>possibly damaging</i>	ja
11. <i>MLL3</i>	NM_170606: Position chr7:151935799: Exon 15: c.T2645C, p.I882T	<i>possibly damaging</i>	ja
12. <i>PSG3</i>	NM_021016: Position chr19:43233329: Exon 5: c.C1189G, p.R397G	<i>possibly damaging</i>	nicht bekannt
13. <i>DCAF7</i>	NM_005828: Position chr17:61660890: Exon 5: c.G556C, p.G186R	<i>possibly damaging</i>	nicht bekannt
14. <i>ECD</i>	NM_001135753: Position chr10:74923593: Exon 2: c.C103G, p.L35V	<i>probably damaging</i>	ja
15. <i>TSPAN14</i>	NM_030927: Position chr10:82271993: Exon 6: c.G544C, p.V182	<i>probably damaging</i>	ja
16. <i>OR51F1</i>	NM_001004752: Position chr11:4791021: Exon 1: c.G127A, p.G43R	<i>probably damaging</i>	ja

Gen	Annotation GRCh37/hg19	PolyPhen-2	Konservierung der Mutationsstelle
17. <i>USP35</i>	NM_020798: Position chr11:77921402: Exon 10: c.G2501C, p.R834P	<i>probably damaging</i>	ja
18. <i>E4F1</i>	NM_004424: Position chr16:2279614: Exon 3: c.C353A, p.A118D	<i>probably damaging</i>	ja
19. <i>ZG16</i>	NM_152338: Position chr16:29791575: Exon 4: c.C339G, p.D113E	<i>probably damaging</i>	ja
20. <i>ZNF223</i>	NM_013361: Position chr19:44571260: Exon 5: c.C1279A, p.P427T	<i>probably damaging</i>	ja
21. <i>SGPP2</i>	NM_152386: Position chr2:223386552: Exon 3: c.G445T, p.V149F	<i>probably damaging</i>	ja
22. <i>BCLAF1</i>	NM_014739: Position chr6:136599391: Exon 4: c.G628A, p.G210S	<i>probably damaging</i>	ja
23. <i>NOTCH4</i>	NM_004557: Position chr6:32170216: Exon 21: c.C3392T:p.P1131L,	<i>probably damaging</i>	ja
24. <i>UBE2J1</i>	NM_016021: Position chr6:90039504: Exon 8: c.A851G, p.H284R	<i>probably damaging</i>	ja
25. <i>OPN1SW</i>	NM_001708: Position chr7:128415071: Exon 2: c.G490A, p.V164I	<i>probably damaging</i>	ja
26. <i>INTS1</i>	NM_001080453: Position chr7:1527546: Exon 19: c.C2366T, p.T789M	<i>probably damaging</i>	ja
27. <i>BAG4</i>	NM_001204878: Position chr8:38065246: Exon 2: c.C487G, p.P163A	<i>probably damaging</i>	ja
28. <i>NEFH</i>	NM_021076: Position chr22:29885498: Exon 4: c.G1869T, p.K623N	<i>probably damaging</i>	nicht vollständig
29. <i>UBTFL1</i>	NM_001143975: Position chr11:89819885: Exon 1: c.A768T, p.R256S	<i>probably damaging</i>	nein
30. <i>CYP2F1</i>	NM_000774: Position chr19:41631417: Exon 9: c.T1172C, p.L391P	<i>probably damaging</i>	nein
31. <i>ZNF432</i>	NM_014650: Position chr19:52537587: Exon 5: c.C1345T, p.R449X	<i>stopgain</i>	ja
32. <i>MYOM3</i>	NM_152372: Position chr1:24417419: Exon 12: c.A1300G, p.I434V	<i>benign</i>	ja
33. <i>OBFC1</i>	NM_024928: Position chr10:105659916: Exon 5: c.G361A, p.E121K	<i>benign</i>	ja
34. <i>C13orf26</i>	NM_152325: Position chr13:31513900: Exon 2: c.T131C, p.V44A	<i>benign</i>	ja
35. <i>CYFIP1</i>	NM_001033028: Position chr15:22956213: Exon 1: c.G157A, p.V53I	<i>benign</i>	ja
36. <i>SULT1A2</i>	NM_001054: Position chr16:28606952: Exon 3: c.G193A, p.G65S	<i>benign</i>	ja
37. <i>C18orf1</i>	NM_001003675: Position chr18:13645512: Exon 3: c.G612A, p.M204I	<i>benign</i>	ja
38. <i>C20orf26</i>	NM_015585: Position chr20:20071504: Exon 7: c.G583A, p.D195N	<i>benign</i>	ja
39. <i>ACAA1</i>	NM_001130410: Position chr3:38175468: Exon 3: c.A298G, p.M100V	<i>benign</i>	ja
40. <i>RAB28</i>	NM_001017979: Position chr4:13378224: Exon 6: c.T518C, p.V173A	<i>benign</i>	ja
41. <i>FRG1</i>	NM_004477: Position chr4:190876301: Exon 5: c.C427A, p.Q143K	<i>benign</i>	ja
42. <i>DNAH5</i>	NM_001369: Position chr5:13754416: Exon 62: c.A10451T, p.H3484L	<i>benign</i>	ja
43. <i>MBLAC2</i>	NM_203406: Position chr5:89770095: Exon 1: c.G15C, p.E5D	<i>benign</i>	ja
44. <i>JARID2</i>	NM_004973: Position chr6:15496915: Exon 7: c.A1459G, p.K487E	<i>benign</i>	ja

Gen	Annotation GRCh37/hg19	PolyPhen-2	Konservierung der Mutationsstelle
45. <i>SEMA3D</i>	NM_152754: Position chr7:84727240: Exon 2: c.T193C, p.S65P	<i>benign</i>	ja
46. <i>TMEM132B</i>	NM_052907: Position chr12:125834111: Exon 2: c.G166A, p.A56T	<i>benign</i>	nicht vollständig
47. <i>LRR1Q1</i>	NM_001079910: Position chr12:85450225: Exon 8: c.A1654C, p.K552Q	<i>benign</i>	nicht vollständig
48. <i>C18orf45</i>	NM_032933: Position chr18:21017812: Exon 1: c.G5A, p.C2Y	<i>benign</i>	nicht vollständig
49. <i>LAMA1</i>	NM_005559: Position chr18:7080048: Exon 3: c.A271G, p.N91D	<i>benign</i>	nicht vollständig
50. <i>KIF15</i>	NM_020242: Position chr3:44879843: Exon 27: c.C3248T, p.S1083L	<i>benign</i>	nicht vollständig
51. <i>WHAMM</i>	NM_001080435: Position chr15:83499472: Exon 9: c.C1763T, p.P588L	<i>benign</i>	nicht vollständig
52. <i>PRDM1</i>	NM_001198: Position chr6:106553383: Exon 5: c.C1348T, p.L450F	<i>benign</i>	nicht vollständig
53. <i>APEH</i>	NM_001640: Position chr3:49720010: Exon 19: c.A1724G, p.H575R	<i>benign</i>	nicht vollständig
54. <i>UNC13D</i>	NM_199242: Position chr17:73827335: Exon 26: c.A2542C, p.I848L	<i>benign</i>	nicht vollständig
55. <i>PTX4</i>	NM_001013658: Position chr16:1536068: Exon 3: c.G1294A, p.G432R	<i>benign</i>	nicht vollständig
56. <i>RGR</i>	NM_001012720: Position chr10:86012708: Exon 4: c.C454A, p.H152N	<i>benign</i>	nicht vollständig
57. <i>ARHGEF19</i>	NM_153213: Position chr1:16525657: Exon 15: c.T2239C, p.W747R	<i>benign</i>	nein
58. <i>OR2L8</i>	NM_001001963: Position chr1:248112754: Exon 1: c.T595C, p.F199L	<i>benign</i>	nein
59. <i>GPR157</i>	NM_024980: Position chr1:9164506: Exon 4: c.G980C, p.G327A	<i>benign</i>	nein
60. <i>DTX2</i>	NM_001102595: Position chr7:76126737: Exon 5: c.C1093T, p.R365C	<i>benign</i>	nein
61. <i>KIR3DL3</i>	NM_153443: Position chr19:55239176: Exon 4: c.G455A, p.R152H	<i>benign</i>	nein
62. <i>CPT1C</i>	NM_001199752: Position chr19:50215114: Exon 17: c.G1915A, p.A639T	<i>benign</i>	nicht bekannt
63. <i>ZNF717</i>	NM_001128223: Position chr3:75788152: Exon 5: c.A622G, p.T208A	<i>benign</i>	nein
64. <i>UBN2</i>	NM_173569: Position chr7:138978680: Exon 17: c.C3947T, p.T1316M	<i>benign</i>	nein
65. <i>PKD1L1</i>	NM_138295: Position chr7:47968953: Exon 7: c.G908A, p.R303Q	<i>benign</i>	nein
66. <i>SSH1</i>	NM_001161331: Position chr12:109186505: Exon 13: c.C1483T, p.P495S	<i>benign</i>	nein
67. <i>MUC17</i>	NM_001040105: Position chr7:100685634: Exon 3: c.C10937T, p.S3646L	<i>benign</i>	nein
68. <i>SLC25A26</i>	NM_173471: Position chr3:66287056: Exon 3: c.G122A, p.S41N	<i>benign</i>	nein
69. <i>AEN</i>	NM_022767: Position chr15:89169738: Exon 2: c.G298A, p.G100R	<i>benign</i>	nein
70. <i>TOX2</i>	NM_032883: Position chr20:42680088: Exon 5: c.G428C, p.S143T	<i>benign</i>	nein
71. <i>MYO1H</i>	NM_001101421: Position chr12:109883374: Exon 30: c.T3002G, p.L1001R	<i>benign</i>	nein
72. <i>MTHFR</i>	NM_005957: Position chr1:11854790: Exon 7: c.C1162T, p.R388C	<i>unknown</i>	ja

Gen	Annotation GRCh37/hg19	PolyPhen-2	Konservierung der Mutationsstelle
73. <i>PINX1</i>	NM_017884: Position chr8:10622935: Exon 7: c.G963C, p.K321N	<i>unknown</i>	ja
74. <i>PLEC</i>	NM_201383: Position chr8:144993053: Exon 32: c.C10948T, p.R3650C	<i>unknown</i>	ja
75. <i>MUC17</i>	NM_001040105: Position chr7:100675940: Exon 3: c.A1243G, p.T415A	<i>unknown</i>	nicht bekannt
76. <i>RECQL5</i>	NM_004259: Position chr17:73624353: Exon 18: c.T2750C, p.L917S	<i>unknown</i>	nein
77. <i>POPDC3</i>	NR_024539: Position chr6:105605893: Exon 4: c.1268_0del	<i>noncoding RNA</i>	ja
78. <i>HCG4</i>	NR_002139: Position chr6:29759680: Exon 1: c.1170_0del	<i>noncoding RNA</i>	ja
79. <i>LOC643837</i>	NR_015368: Position chr1:789515: Exon 7: c.1318delA	<i>noncoding RNA</i>	ja
80. <i>SLC7A5P2</i>	NR_002594: Position chr16:21531464: Exon 1: c.G302A	<i>noncoding RNA</i>	ja
81. <i>TRIB2</i>	NR_027303: Position chr2:12882826: Exon 3: c.2807_0del	<i>noncoding RNA</i>	ja
82. <i>LOC96610</i>	NR_027293: Position chr22:22664216: Exon 7: c.G853A	<i>noncoding RNA</i>	ja
83. <i>LOC285768</i>	NR_027116: Position chr6:1100888: Exon 1: c.A680T	<i>noncoding RNA</i>	ja
84. <i>LINC00174</i>	NR_026873: Position chr7:65841880: Exon 5: c.C3575T	<i>noncoding RNA</i>	ja
85. <i>ZNF323</i>	NR_024164: Position chr6:28293409: Exon 3: c.1679_1680insG	<i>noncoding RNA</i>	nicht vollständig
86. <i>KCNIP4-IT1</i>	NR_002813: Position chr4:21853830: Exon 1: c.982delG	<i>noncoding RNA</i>	nicht vollständig
87. <i>ANKRD20A8P</i>	NR_003366: Position chr2:95488771: Exon 10: c.G947A	<i>noncoding RNA</i>	nicht vollständig
88. <i>CROCCP2</i>	NR_026752: Position chr1:16945452: Exon 7: c.G1064A	<i>noncoding RNA</i>	nein
89. <i>PWRN2</i>	NR_026647: Position chr15:24413033: Exon 1: c.C2021T	<i>noncoding RNA</i>	nein
90. <i>LOC127841</i>	NR_027022: Position chr1:204338127: Exon 1: c.A721C	<i>noncoding RNA</i>	nein
91. <i>LOC149134</i>	NR_015422: Position chr1:246954534: Exon 2: c.A869C	<i>noncoding RNA</i>	nicht bekannt
92. <i>ALS2CL</i>	NR_033815: Position chr3:46730942: Exon 2: c.C76T	<i>noncoding RNA</i>	nein
93. <i>DNM1P46</i>	NR_003260: Position chr15:100331641: Exon 5: c.T2551C	<i>noncoding RNA</i>	nicht bekannt
94. <i>C21orf96</i>	NR_026812: Position chr21:36410912: Exon 1: c.C812G	<i>noncoding RNA</i>	nicht bekannt
95. <i>SRPK1</i>	NR_034069: Position chr6:35801739: Exon 16: c.3459_0del	<i>noncoding RNA</i>	nicht bekannt
96. <i>SNORD114-2</i>	NR_003194: Position chr14:101418193: Exon 1: c.G1T	<i>noncoding RNA</i>	nicht bekannt
97. <i>GOLGA8IP</i>	NR_024074: Position chr15:23261024: Exon 9: c.A998T	<i>noncoding RNA</i>	nicht bekannt
98. <i>LINC00482</i>	NR_038080: Position chr17:79278528: Exon 4: c.G1066A	<i>noncoding RNA</i>	nicht bekannt
99. <i>LOC146336</i>	NR_027242: Position chr16:1115495: Exon 4: c.1520_1521insA	<i>noncoding RNA</i>	nicht bekannt
100. <i>ANKRD30BL</i>	NR_027020: Position chr2:133015302: Exon 1: c.A241C	<i>noncoding RNA</i>	nicht bekannt



Gen	Annotation GRCh37/hg19	PolyPhen-2	Konservierung der Mutationsstelle
101. <i>FXVD6-FXYD2</i>	NR_037907: Position chr11:117747578: Exon 1: c.C169A	<i>noncoding RNA</i>	nein
102. <i>FAM99A</i>	NR_026643: Position chr11:1688904: Exon 3: c.1248_0del	<i>noncoding RNA</i>	nein
103. <i>PAR5</i>	NR_022008: Position chr15:25230881: Exon 1: c.A875T	<i>noncoding RNA</i>	nein
104. <i>FLJ12825</i>	NR_026655: Position chr12:54515621: Exon 2: c.T3545C	<i>noncoding RNA</i>	nein

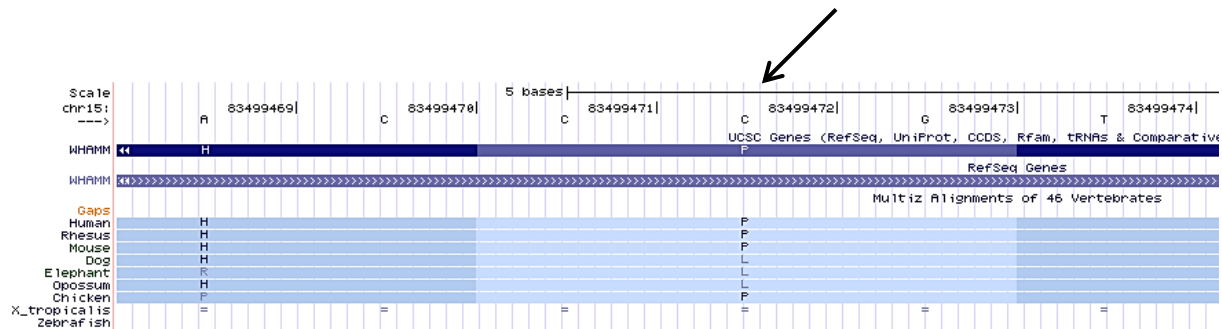
Exemplarisch wurde für die Varianten in den Genen *GJB4* (1.), *ZG16* (19.), *INTS1* (26.), *CYFIP1* (35.), *WHAMM* (51.) und *PINX1* (73.) der Tabelle 29 die Ko-Segregation in der Familie mittels Sanger-Sequenzierung überprüft (Tabelle 30).

**Tabelle 30: Ko-Segregation von sechs Varianten in der DCM-Familie 009-1229**

ID	Phänotyp *	<i>GJB4</i> - Deletion	<i>ZG16</i> - Variante	<i>INTS1</i> - Variante	<i>CYFIP1</i> - Variante	<i>WHAMM</i> - Variante	<i>PINX1</i> - Variante
01	DCM, 68 Jahre	positiv	positiv	positiv	positiv	positiv	positiv
02	DCM, 41 Jahre	positiv	positiv	positiv	positiv	positiv	positiv
03	DCM, 16 Jahre	positiv	positiv	positiv	positiv	positiv	positiv
04	Nicht betroffen, 12 Jahre	negativ	positiv	positiv	negativ	negativ	negativ
05	Nicht betroffen, 44 Jahre	positiv	negativ	negativ	positiv	negativ	positiv

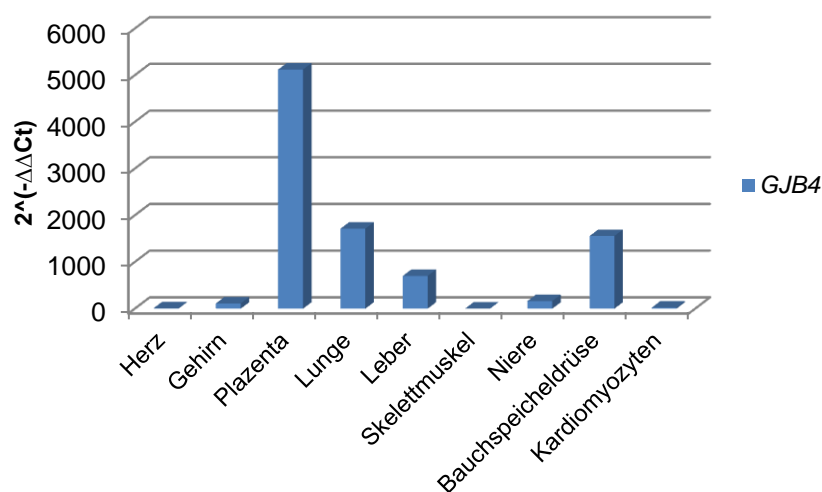
\* Alter bei Studien-Einschluss bei betroffen und nicht betroffen

Die Sequenzierung ergab, dass nur die Variante im Gen *WHAMM* eine vollständige Ko-Segregation mit dem Phänotyp zeigte. Die nicht erkrankten Familienmitglieder (ID 04/ID 05) wiesen gegenüber den DCM erkrankten Familienmitgliedern (ID 01/ID 02/ID 03) die Variante nicht auf (Tabelle 30). Laut einer PolyPhen-2 Vorhersage wurde diese Variante als nicht schädlich für das entstehende Protein eingestuft (Tabelle 29). Weiterhin zeigte sich bei Betrachtung des Konservierungsgrades der Mutationsstelle, dass nur in manchen Organismen (Mensch, Affe, Maus, Huhn) die Aminosäure Prolin erhalten geblieben war. Die anderen Organismen wiesen an der Stelle die Aminosäure Leucin auf. Durch die Variante im Gen *WHAMM* wird ebenso die Aminosäure Leucin in das Protein eingebaut. Aus diesem Grund wurde die Konservierung der Mutationsstelle als nicht vollständig eingestuft (Abbildung 29/Tabelle 29).



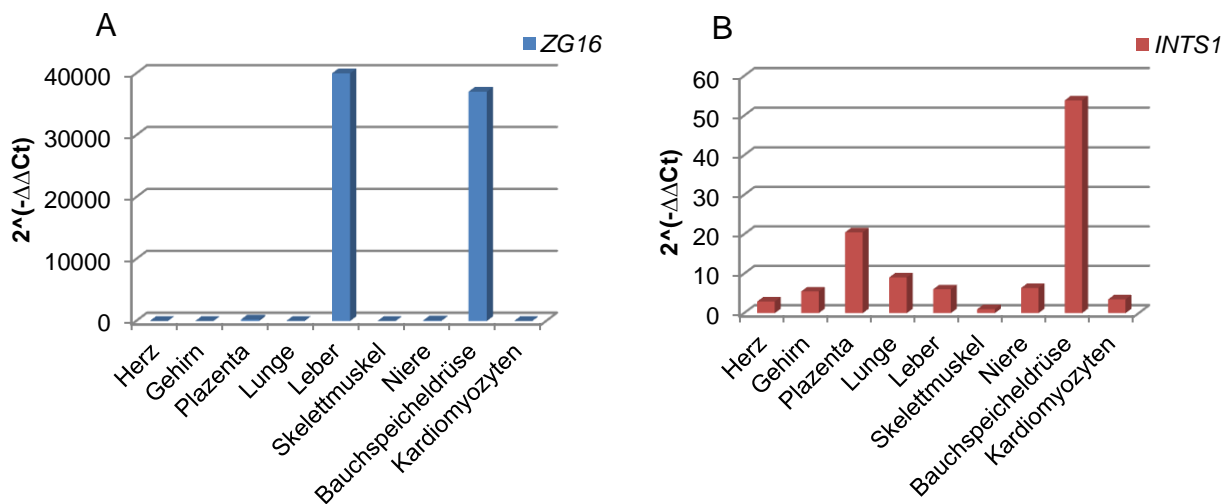
**Abbildung 29: Konservierung der Mutationsstelle der *WHAMM*-Variante.** Der Pfeil markiert die Position der Mutationsstelle im Gen *WHAMM*. Die Konservierung ist für folgende Organismen gezeigt: Mensch (*human*), Rhesusaffe (*rhesus*), Maus (*mouse*), Hund (*dog*), Elefant (*elephant*), Opossum (*opossum*) und Huhn (*chicken*) (<http://genome.ucsc.edu/>).

Um die Kausalität der Varianten neben der Ko-Segregation zu überprüfen, wurde die Expression der Gene in verschiedenen Geweben/Zellen untersucht (vgl. 3.7.3 und 3.7.4). Exemplarisch ist die relative Expression von den Genen *GJB4*, *ZG16* und *INTS1* dargestellt. Ein starker Effekt auf das Protein durch die Veränderung des Leserahmens wurde bei der Deletion im Gen *GJB4* vermutet (Tabelle 29). Für das Gen wurde eine hohe Expression in Plazenta, Lunge, Leber und Bauchspeicheldrüse beobachtet. In Herz und Kardiomyozyten wurde keine oder nur schwache Expression gemessen (Abbildung 30).



**Abbildung 30: Relative Expression von *GJB4*.** Als *housekeeper* wurde *GAPDH* verwendet, die Expression im Skelettmuskel diente als Kalibrator. Auf der Abszisse sind verschiedene Gewebe/Zellen aufgetragen, auf der Ordinate der  $2^{-\Delta\Delta Ct}$ -Wert.

Die Varianten in den Genen *ZG16* und *INTS1* hatten laut PolyPhen-2 einen schädlichen Einfluss auf die entstehenden Proteine (Tabelle 29). *ZG16* zeigte eine sehr hohe Expression in Leber und Bauchspeicheldrüse, allerdings keine Expression in Herz oder Kardiomyozyten (Abbildung 31A). Die Expression von *INTS1* wurde vor allem in der Bauchspeicheldrüse nachgewiesen. Zudem zeigte sich eine ubiquitäre Expression von *INTS1* in den Geweben, auch in Herz und Kardiomyozyten (Abbildung 31B).



**Abbildung 31: Relative Expression von *ZG16* (A) und *INTS1* (B). Als *housekeeper* wurde *GAPDH* verwendet, die Expression im Skelettmuskel diente als Kalibrator. Auf der Abszisse sind verschiedene Gewebe/Zellen aufgetragen, auf der Ordinate der  $2^{-(\Delta\Delta Ct)}$ -Wert.**

#### 4.5.2. Zusätzlich identifizierte Varianten der nicht modifizierten Proben

Zusätzlich zu den 104 Varianten wurden bei den Proben mit dem nicht modifizierten Protokoll 42 Varianten von ID 01 und ID 03 geteilt (Tabelle 31). Für elf Varianten wurde durch PolyPhen-2 oder aufgrund einer vorliegenden Insertion/Deletion ein starker Effekt auf die Proteine vorhergesagt. Außerdem wurden zehn weitere SNVs im kodierenden Bereich detektiert und 21 SNVs in nicht-kodierenden Regionen. Für viele Varianten ( $n=11$ ) in nicht-kodierenden RNAs konnte keine Information bezüglich der Konservierung ermittelt werden (Tabelle 31).

**Tabelle 31: Zusätzlich geteilte Varianten der Exome von DCM-Familie 009-1229 (nicht modifiziert)**

Gen	Annotation GRCh37/hg19	PolyPhen-2	Konservierung der Mutationsstelle
1. <i>GPSM1</i>	NM_015597: Position chr9:139235485: Exon 9: c.1242delC, p.P414fs	<i>frameshift</i>	ja
2. <i>ZNF598</i>	NM_178167: Position chr16:2049898: Exon 9: c.1654_1655insAGG, p.R552delinsRR	<i>nonframeshift</i>	nein
3. <i>CSF2RB</i>	NM_000395: Position chr22:37319327: Exon 3: c.A118C, p.T40P	<i>possibly damaging</i>	nicht vollständig
4. <i>HSPG2</i>	NM_005529: Position chr1:22186113: Exon 42: c.A5239C, p.T1747P	<i>probably damaging</i>	ja
5. <i>HSPG2</i>	NM_005529: Position chr1:22191361: Exon 36: c.C4601G, p.P1534R	<i>probably damaging</i>	ja
6. <i>P2RY4</i>	NM_002565: Position chrX:69478801: Exon 1: c.G674C, p.R225P	<i>probably damaging</i>	ja
7. <i>NPPA</i>	NM_006172: Position chr1:11907414: Exon 2: c.A206C, p.E69A	<i>probably damaging</i>	ja
8. <i>DNAH5</i>	NM_001369: Position chr5:13721137: Exon 71: c.G12251A, p.R4084Q	<i>probably damaging</i>	ja
9. <i>GPR98</i>	NR_003149: Position chr5:89943581: Exon 17: c.G3385A, p.G1129S	<i>probably damaging</i>	ja
10. <i>NLRC3</i>	NM_178844: Position chr16:3602228: Exon 10: c.C2320T, p.L774F	<i>probably damaging</i>	nicht vollständig
11. <i>SLFNL1</i>	NM_144990: Position chr1:41486083: Exon 2: c.G250T, p.E84X	<i>stopgain</i>	ja
12. <i>FLNB</i>	NM_001164317: Position chr3:58145363: Exon 43: c.A7064C, p.H2355P	<i>benign</i>	ja
13. <i>SPTBN4</i>	NM_020971: Position chr19:41062980: Exon 26: c.C5341G, p.R1781G	<i>benign</i>	ja
14. <i>RBMX</i>	NR_028476: Position chrX:135957501: Exon 6: c.A736C, p.I246L	<i>benign</i>	ja
15. <i>UBE2Q1</i>	NM_017582: Position chr1:154528361: Exon 2: c.T407G, p.V136G	<i>benign</i>	nicht vollständig
16. <i>UNC13D</i>	NM_199242: Position chr17:73825036: Exon 31: c.G2983C, p.A995P	<i>benign</i>	nicht vollständig
17. <i>SEZ6L</i>	NM_001184776: Position chr22:26709766: Exon 9: c.T1913G, p.V638G	<i>unknown</i>	ja
18. <i>COL4A</i>	NM_033641: Position chrX:107400382: Exon 45: c.A4921C, p.T1641P	<i>unknown</i>	ja
19. <i>ZNF595</i>	NM_182524: Position chr4:86018: Exon 4: c.T623C, p.L208P	<i>unknown</i>	ja
20. <i>GAS6</i>	NM_000820: Position chr13:114537583: Exon 8: c.A775C, p.T259P	<i>unknown</i>	nicht vollständig
21. <i>MUC16</i>	NM_024690: Position chr19:8999446: Exon 56: c.A40729G, p.S13577G	<i>unknown</i>	nein
22. <i>ACBD5</i>	NR_024150: Position chr10:27485480: Exon 12: c.A2402G	<i>noncoding RNA</i>	ja
23. <i>CROCCP2</i>	NR_026752: Position chr1:16946437: Exon 6: c.G820A	<i>noncoding RNA</i>	ja
24. <i>ANKRD20A9P</i>	NR_027995: Position chr13:19414116: Exon 1: c.G31994A	<i>noncoding RNA</i>	ja
25. <i>CREBZF</i>	NR_028026: Position chr11:85368940: Exon 5: c.A4286G	<i>noncoding RNA</i>	nicht vollständig
26. <i>ANKRD20A9P</i>	NR_027995: Position chr13:19413822: Exon 1: c.32288_32289insC	<i>noncoding RNA</i>	nein
27. <i>FOXP2</i>	NR_033766: Position chr7:114330203: Exon 20: c.G2755A	<i>noncoding RNA</i>	nein

Gen	Annotation GRCh37/hg19	PolyPhen-2	Konservierung der Mutationsstelle
28. <i>UGGT1</i>	NR_027671: Position chr2:128949498: Exon 41: c.C7181T	<i>noncoding RNA</i>	nein
29. <i>ASB13</i>	NR_024581: Position chr10:5681690: Exon 5: c.A1706G	<i>noncoding RNA</i>	nein
30. <i>MST1P9</i>	NR_002729: Position chr1:17086087: Exon 7: c.810_811insG	<i>noncoding RNA</i>	nicht bekannt
31. <i>LOC399744</i>	NR_024497: Position chr10:38738912: Exon 8: c.G3084T	<i>noncoding RNA</i>	nicht bekannt
32. <i>FAM66C</i>	NR_026788: Position chr12:8349048: Exon 7: c.T1907C	<i>noncoding RNA</i>	nicht bekannt
33. <i>LOC729082</i>	NR_026757: Position chr15:41591334: Exon 4: c.1433_0del	<i>noncoding RNA</i>	nicht bekannt
34. <i>RPL23AP64</i>	NR_003040: Position chr11:118873850: Exon 1: c.G454A	<i>noncoding RNA</i>	nicht bekannt
35. <i>FAM182A</i>	NR_026713: Position chr20:26064303: Exon 5: c.G1861C	<i>noncoding RNA</i>	nicht bekannt
36. <i>LOC388692</i>	NR_027002: Position chr1:149291003: Exon 1: c.C11528T	<i>noncoding RNA</i>	nicht bekannt
37. <i>LINC00294</i>	NR_015451: Position chr11:33098288: Exon 1: c.G593T	<i>noncoding RNA</i>	nicht bekannt
38. <i>GIMAP6</i>	NR_024115: Position chr7:150324547: Exon 3: c.A1339G	<i>noncoding RNA</i>	nicht bekannt
39. <i>MST1P9</i>	NR_002729: Position chr1:17086003: Exon 7: c.891_0del	<i>noncoding RNA</i>	nicht bekannt
40. <i>FAM182B</i>	NR_026714: Position chr20:25754033: Exon 4: c.T1682C	<i>noncoding RNA</i>	nicht bekannt
41. <i>LOC150622</i>	NR_026832: Position chr2:6108879: Exon 2: c.A397T	<i>noncoding RNA</i>	nein
42. <i>ANKRD20A8P</i>	NR_003366: Position chr2:95463733: Exon 18: c.3063_0del	<i>noncoding RNA</i>	nein

Von den 42 Varianten wurde die Genregion von drei Varianten mittels Sanger-Sequenzierung überprüft und die Ko-Segregation in der Familie betrachtet (Tabelle 32). Es handelte sich um die Varianten in den Genen *ZNF598* (2.), *NPPA* (7.) und *GPR98* (9.) der Tabelle 31. Es zeigte sich keine vollständige Ko-Segregation mit dem Phänotyp. Alle drei Varianten wurden in ID 05, einer nicht betroffenen Person (Alter bei Studien-Einschluss: 44 Jahre), identifiziert. (Tabelle 32).

**Tabelle 32: Ko-Segregation von drei Varianten in der DCM-Familie 009-1229**

ID	Phänotyp *	<i>ZNF598</i> -Insertion	<i>NPPA</i> -Variante	<i>GPR98</i> -Variante
01	DCM, 68 Jahre	positiv	positiv	positiv
02	DCM, 41 Jahre	negativ	positiv	positiv
03	DCM, 16 Jahre	positiv	positiv	positiv
04	Nicht betroffen, 12 Jahre	positiv	negativ	negativ
05	Nicht betroffen, 44 Jahre	positiv	positiv	positiv

\* Alter bei Studien-Einschluss bei betroffen und nicht betroffen

### 4.5.3. Zusätzlich identifizierte Varianten der modifizierten Proben

Bei den modifizierten Proben wurden 55 Varianten identifiziert, die bei den nicht modifizierten Proben nicht detektiert wurden (Tabelle 33). Dabei handelte es sich im kodierenden Bereich um fünf Indels und 37 SNVs. In nicht-kodierenden Regionen wurden 13 Varianten identifiziert. Von den 37 SNVs der kodierenden Region wurden mit PolyPhen-2 23 Varianten als möglich oder wahrscheinlich schädlich für die entstehenden Proteine eingestuft. Eine Variante im Gen *ZW10* (29.) führte zur Einführung eines Stoppkodons (Tabelle 33).

**Tabelle 33: Zusätzlich geteilte Varianten der Exome von DCM-Familie 009-1229 (modifiziert)**

Gen	Annotation GRCh37/hg19	PolyPhen-2	Konservierung der Mutationsstelle
1. <i>TDG</i>	NM_003211: Position chr12:104373734: Exon 3: c.292_293insA, p.I98fs	<i>frameshift</i>	ja
2. <i>HNRNPK</i>	NM_002140: Position chr9:86587833: Exon 10: c.571_572insA, p.R191fs	<i>frameshift</i>	ja
3. <i>RBMX</i>	NM_002139: Position chrX:135960147: Exon 4: c.315_316insTT, p.P105fs	<i>frameshift</i>	nicht vollständig
4. <i>TRIM52</i>	NM_032765: Position chr5:180687440: Exon 1: c.373_375del, p.125_125del	<i>nonframeshift</i>	ja
5. <i>ASPHD1</i>	NM_181718: Position chr16:29912807: Exon 1: c.515_516insTGG, p.G172delinsGG	<i>nonframeshift</i>	nicht vollständig
6. <i>TEKT4</i>	NM_144705: Position chr2:95542416: Exon 6: c.A1210G, p.I404V	<i>possibly damaging</i>	ja
7. <i>USP26</i>	NM_031907: Position chrX:132162006: Exon 1: c.G243T, p.L81F	<i>possibly damaging</i>	ja
8. <i>FAM163A</i>	NM_173509: Position chr1:179783095: Exon 5: c.C275G, p.A92G	<i>possibly damaging</i>	nicht vollständig
9. <i>DMWD</i>	NM_004943: Position chr19:46289009: Exon 3: c.C1745T, p.A582V	<i>possibly damaging</i>	nicht vollständig
10. <i>C13orf35</i>	NM_207440: Position chr13:113333878: Exon 2: c.A185C, p.D62A	<i>possibly damaging</i>	nein
11. <i>DLGAP3</i>	NM_001080418: Position chr1:35370204: Exon 1: c.T781G, p.W261G	<i>probably damaging</i>	ja
12. <i>MACF1</i>	NM_012090: Position chr1:39945661: Exon 93: c.A15886C, p.T5296	<i>probably damaging</i>	ja
13. <i>CD2BP2</i>	NM_006110: Position chr16:30364976: Exon 5: c.C521G, p.A174G	<i>probably damaging</i>	ja
14. <i>CNTNAP2</i>	NM_014141: Position chr7:147183121: Exon 11: c.A1765C, p.T589P	<i>probably damaging</i>	ja
15. <i>C14orf49</i>	NM_152592: Position chr14:95910923: Exon 9: c.C1675A, p.P559T	<i>probably damaging</i>	ja
16. <i>DSG2</i>	NM_001943: Position chr18:29125917: Exon 15: c.A2568C, p.K856N	<i>probably damaging</i>	ja
17. <i>WDR52</i>	NM_001164496: Position chr3:113063479: Exon 23: c.T3146G, p.V1049G	<i>probably damaging</i>	ja
18. <i>PIF1</i>	NM_025049: Position chr15:65114504: Exon 4: c.G778C, p.A260P	<i>probably damaging</i>	ja
19. <i>AARS2</i>	NM_020745: Position chr6:44280885: Exon 1: c.T176G, p.V59G	<i>probably damaging</i>	ja

Gen	Annotation GRCh37/hg19	PolyPhen-2	Konservierung der Mutationsstelle
20. <i>IFI30</i>	NM_006332: Position chr19:18286142: Exon 3: c.T337G, p.W113G	<i>probably damaging</i>	ja
21. <i>TRIM14</i>	NM_014788: Position chr9:100849907: Exon 6: c.G1174C, p.G392R	<i>probably damaging</i>	ja
22. <i>FRMD4A</i>	NM_018027: Position chr10:13699338: Exon 22: c.A2251C, p.T751P	<i>probably damaging</i>	ja
23. <i>NAV2</i>	NM_001111018: Position chr11:20101690: Exon 25: c.A5068C, p.T1690P	<i>probably damaging</i>	ja
24. <i>TDG</i>	NM_003211: Position chr12:104373677: Exon 3: c.C235A, p.P79T	<i>probably damaging</i>	ja
25. <i>CCDC157</i>	NM_001017437: Position chr22:30762233: Exon 3: c.G244A; p.D82N	<i>probably damaging</i>	ja
26. <i>C6orf138</i>	NM_207499: Position chr6:48036327: Exon 1: c.T65G, p.V22G	<i>probably damaging</i>	nicht vollständig
27. <i>CACNA1I</i>	NM_021096: Position chr22:40066873: Exon 26: c.A4453C, p.T1485P	<i>probably damaging</i>	nicht vollständig
28. <i>MAGEA11</i>	NM_001011544: Position chrX:148798267: Exon 5: c.T1034G, p.V345G	<i>probably damaging</i>	nein
29. <i>ZW10</i>	NM_004724: Position chr11:113614711: Exon 10: c.G1324T, p.E442X	<i>stopgain</i>	nicht vollständig
30. <i>GPR133</i>	NM_198827: Position chr12:131456080: Exon 4: c.T265G, p.Y89D	<i>benign</i>	ja
31. <i>TBX21</i>	NM_013351: Position chr17:45820022: Exon 2: c.A538C, p.T180P	<i>benign</i>	ja
32. <i>LRRC4</i>	NM_022143: Position chr7:127670327: Exon 2: c.G367C, p.A123P	<i>benign</i>	ja
33. <i>AGBL5</i>	NM_021831: Position chr2:27292989: Exon 15: c.G2519A, p.R840H	<i>benign</i>	nicht vollständig
34. <i>DIDO1</i>	NM_001193370: Position chr20:61528199: Exon 7: c.A1738C, p.T580P	<i>benign</i>	nein
35. <i>FAM186A</i>	NM_001145475: Position chr12:50745703: Exon 4: c.A4912C, p.T1638P	<i>benign</i>	nein
36. <i>ZNF98</i>	NM_001098626: Position chr19:22575777: Exon 4: c.A260G, p.Y87C	<i>benign</i>	nein
37. <i>ODZ2</i>	NM_001122679: Position chr5:167642193: Exon 21: c.A3967G, p.T1323A	<i>unknown</i>	ja
38. <i>HLA-C</i>	NM_002117: Position chr6:31239613: Exon 2: c.C106A, p.R36S	<i>unknown</i>	ja
39. <i>DNAH8</i>	NM_001206927: Position chr6:38810444: Exon 35: c.G4610A, p.R1537H	<i>unknown</i>	ja
40. <i>SLC45A4</i>	NM_001080431: Position chr8:142221750: Exon 8: c.T2188G, p.C730G	<i>unknown</i>	nicht vollständig
41. <i>LILRB5</i>	NM_001081442: Position chr19:54760381: Exon 3: c.C326T, p.P109L	<i>unknown</i>	nicht vollständig
42. <i>WNK1</i>	NM_001184985: Position chr12:993829: Exon 19: c.G4639T, p.A1547S	<i>unknown</i>	nein
43. <i>GK3P</i>	NR_026575: Position chr4:166201039: Exon 1: c.A137G	<i>noncoding RNA</i>	ja
44. <i>FRG1B</i>	NR_003579: Position chr20:29628328: Exon 5: c.C630G	<i>noncoding RNA</i>	ja
45. <i>AK2</i>	NR_037591: Position chr1:33475556: Exon 8: c.T1774G	<i>noncoding RNA</i>	ja
46. <i>LOC441666</i>	NR_024380: Position chr10:42833154: Exon 3: c.A821G	<i>noncoding RNA</i>	ja
47. <i>LOC100507557</i>	NR_038244: Position chr6:146206203: Exon 4: c.A2585G	<i>noncoding RNA</i>	nicht vollständig

Gen	Annotation GRCh37/hg19	PolyPhen-2	Konservierung der Mutationsstelle
48. <i>TTC31</i>	NR_027749: Position chr2:74721679: Exon 12: c.2808delG	<i>noncoding RNA</i>	nicht vollständig
49. <i>LOC283050</i>	NR_024431: Position chr10:80703432: Exon 10: c.A2084G	<i>noncoding RNA</i>	nicht vollständig
50. <i>LOC441666</i>	NR_024380: Position chr10:42830874: Exon 3: c.C3101T	<i>noncoding RNA</i>	nicht bekannt
51. <i>RPL23AP64</i>	NR_003040: Position chr11:118873865: Exon 1: c.T439A	<i>noncoding RNA</i>	nicht bekannt
52. <i>FAM182A</i>	NR_026713: Position chr20:26063993: Exon 5: c.A1551G	<i>noncoding RNA</i>	nicht bekannt
53. <i>RNF126P1</i>	NR_002818: Position chr17:55123871: Exon 1: c.1033_1034insGACT	<i>noncoding RNA</i>	nicht bekannt
54. <i>TCL6</i>	NR_028288: Position chr14:96117918: Exon 1: c.T404G	<i>noncoding RNA</i>	nicht bekannt
55. <i>ANKRD20A9P</i>	NR_027995: Position chr13:19436209: Exon 1: c.C9901T	<i>noncoding RNA</i>	nein

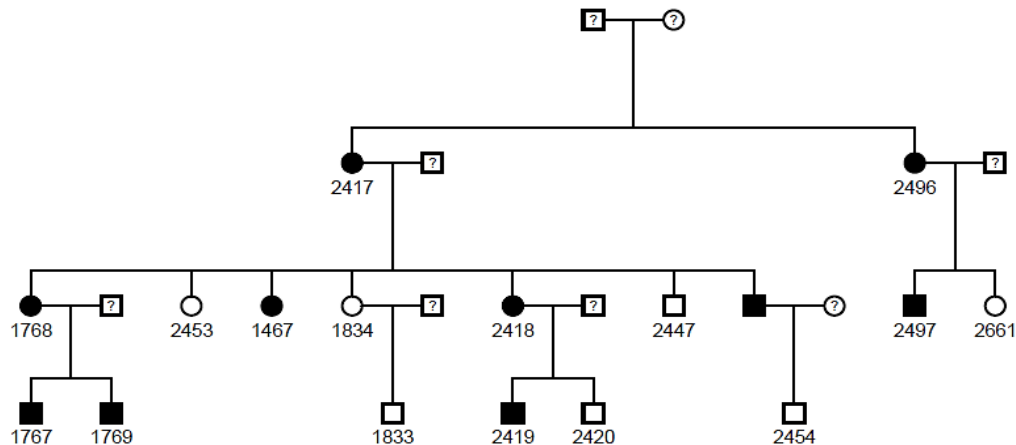
Von den Varianten der Tabelle 33 wurde exemplarisch die *ZNF98*-Genregion (36.) sequenziert. Die Sequenzierung zeigte, dass es sich um eine falsch positive Variante handelte. In ID 01<sub>modif.</sub> und ID 03<sub>modif.</sub> konnte die ursprünglich durch die Exom-Sequenzierung detektierte Variante nicht wieder gefunden werden.

Zusammenfassend wurden in den zwei Exom-Sequenzen 201 Varianten identifiziert, die kausal für die Entstehung von DCM sein konnten. Bislang wurde eine Variante im Gen *ZNF98* als falsch positive Variante identifiziert und neun weitere Varianten bestätigt.



#### 4.6. Exom-Sequenzierung HCM-Familie 757

Bei der HCM-Familie 757 waren zehn Individuen über drei Generationen an HCM erkrankt (Abbildung 32). Die klinischen Charakteristika der Familie sind in Tabelle 43 unter 8.1 zusammengefasst.



**Abbildung 32: Stammbaum der HCM-Familie 757. Quadrat: Mann, Kreis: Frau, ausgefüllte Symbole: betroffen, offene Symbole: nicht betroffen, durchgestrichene Symbole: verstorben, ?: Phänotyp unbekannt.**

Drei betroffene Familienmitglieder (ID 1769, ID 2419, ID 2497) wurden für die Exom-Sequenzierung ausgewählt. Die Proben wurden nur mit dem nicht modifizierten Protokoll vorbereitet. Es zeigte sich, dass die generierten Exom-Sequenzen ähnliche Eigenschaften aufwiesen. Bei allen Proben wurden circa 12,4% der *reads* als Duplikate entfernt. Von den verbleibenden *reads* wurden >98% zugeordnet, was 4,5 Gb – 4,8 Gb Sequenz entsprach. 55% – 57% dieser Sequenzen waren *on target*. Dabei erreichten 96% aller Basen eine *coverage* von  $\geq 1x$  und 64% - 68% eine *coverage* von  $\geq 20x$  (Tabelle 34).

**Tabelle 34: Information der generierten Exom-Sequenzen für HCM-Familie 757**

ID	Duplikate in %	<i>reads/mapped reads</i>	<i>mapped reads in %</i>	<i>mapped sequence in bp</i>	<i>on target in %</i>	<i>coverage <math>\geq 1x</math> in %</i>	<i>coverage <math>\geq 4x</math> in %</i>	<i>coverage <math>\geq 8x</math> in %</i>	<i>coverage <math>\geq 20x</math> in %</i>
1769	12,3	65.229.278/ 64.439.414	98,8	4.832.956.050	57	96	92	86	68
2419	12,5	60.833.286/ 60.132.910	98,8	4.509.968.250	57	96	91	86	65
2497	12,4	61.114.084/ 60.345.851	98,7	4.525.938.825	55	96	92	86	64

Pro Exom wurden die Varianten detektiert und annotiert (vgl. 8.6 Tabelle 55). Die Anzahl der ausgewählten Varianten (heterozygote nicht-synonyme Varianten, Indels und Spleiß-Varianten ohne rs-Nummer) ist in Tabelle 35 dargestellt. Für ID 1769 wurden 717 Varianten ausgewählt. Für ID 2419 kamen 671 Varianten in Frage und für ID 2497 wurden 676 Varianten weiter betrachtet (Tabelle 35).

**Tabelle 35: Anzahl detektierter Varianten der HCM-Familie 757**

<b>ID</b>	<b>Anzahl heterozygoter nicht-synonymer Varianten, Indels und Spleiß-Varianten ohne rs-Nummer</b>
1769	717
2419	671
2497	676

#### 4.6.1. Identifizierte Varianten

Der Vergleich dieser Varianten zwischen den drei Familienmitgliedern ergab 55 geteilte Varianten. 35 dieser Varianten wurden als paralog eingestuft und gefiltert. Vertreter der Gen-Familie *ANKRD* und *FAM* waren dabei mehrheitlich vertreten (Tabelle 36).

**Tabelle 36: Paraloge Genregionen der HCM-Familie 757**

Position	Gen	Paraloge Gene
chr1:145302676	<i>NBPF10</i>	<i>NBPF15</i> , <i>NBFP16</i>
chr1:148004625	<i>NBPF14</i>	<i>NBPF15</i> , <i>NBFP16</i>
chr2:114356500	<i>WASH2P</i>	<i>WASH3P</i>
chr2:132905792	<i>ANKRD30BL</i>	<i>ANKRD30BP2</i>
chr4:144801662	<i>GYPE</i>	<i>GYPB</i>
chr8:11985979	<i>LOC392196</i>	<i>FAM66B</i>
chr8:12040786	<i>FAM86B1</i>	<i>FAM86B2</i>
chr8:7629232	<i>FAM90A10</i>	<i>FAM90A7</i>
chr9:97080947	<i>FAM22F</i>	<i>FAM22G</i>
chr10:135440080	<i>FRG2B</i>	<i>FRG2</i>
chr10:49239525	<i>LOC399753</i>	<i>AGAP9</i>
chr10:51748530	<i>AGAP6</i>	<i>AGAP8</i>
chr10:51768676	<i>AGAP6</i>	<i>AGAP8</i>
chr12:8353318	<i>FAM66C</i>	<i>FAM66B</i>
chr13:19409574	<i>ANKRD20A9P</i>	<i>ANKRD20A5P</i>
chr13:19412955	<i>ANKRD20A9P</i>	<i>ANKRD20A5P</i>
chr13:19413087	<i>ANKRD20A9P</i>	<i>ANKRD20A5P</i>
chr13:19999130	<i>TPTE2</i>	<i>TPTE2P1</i>
chr15:20771946	<i>GOLGA8C</i>	<i>GOLGA8G</i> , <i>GOLGA8F</i>
chr15:89399986	<i>ACAN</i>	<i>ACAN</i>
chr16:22545897	<i>LOC100132247</i>	<i>LOC613037</i>
chr18:14542791	<i>POTEC</i>	<i>POTEB</i>
chr20:25753403	<i>FAM182B</i>	<i>FAM182A</i>
chr20:25753571	<i>FAM182B</i>	<i>FAM182A</i>
chr20:25754505	<i>FAM182B</i>	<i>FAM182A</i>
chr20:26064330	<i>FAM182A</i>	<i>FAM182B</i>
chr20:26067223	<i>FAM182A</i>	<i>FAM182B</i>
chr20:26067232	<i>FAM182A</i>	<i>FAM182B</i>
chr21:9908223	<i>TEKT4P2</i>	<i>MAFIP</i>
chr22:21044704	<i>POM121L4P</i>	<i>POM121L10P</i>
chr22:21044717	<i>POM121L4P</i>	<i>POM121L10P</i>
chr22:22974921	<i>POM121L1P</i>	<i>POM121L10P</i>
chr22:23915620	<i>IGLL1</i>	<i>IGLL3P</i>
chr22:25024072	<i>GGT1</i>	<i>GGTLC1</i>
chr22:25043065	<i>POM121L10P</i>	<i>POM121L1P</i>

Somit ergab die Auswertung 20 geteilte Varianten. Neben zwei Deletionen in den Genen *MRO* und *RYR3* wurden zehn SNVs in kodierenden Regionen und acht Varianten in nicht-kodierenden RNAs identifiziert (Tabelle 37). Sieben SNVs wurden durch PolyPhen-2 als schädlich für die entstehenden Proteine vorhergesagt. Die Konservierung dieser Varianten war oft gegeben, nur die Mutationsstellen in den Genen *SPATC1* und *TTC24* waren nicht vollständig konserviert (Tabelle 37).

Tabelle 37: Geteilte Varianten der Exome von HCM-Familie 757

Gen	Annotation GRCh37/hg19	PolyPhen-2	Konservierung der Mutationsstelle
1. <i>MRO</i>	NM_001127176: Position chr18:48331614: Exon 4: c.380_381del, p.127_127del	<i>frameshift</i>	nicht vollständig
2. <i>RYR3</i>	NM_001036: Position chr15:34137086: Exon 93: c.13320_13322del, p.4440_4441del	<i>nonframeshift</i>	nicht vollständig
3. <i>FAM104B</i>	NM_001166703: Position chrX:55185562: Exon 2: c.G117T, p.E39D	<i>possibly damaging</i>	ja
4. <i>SPATC1</i>	NM_198572: Position chr8:145086736: Exon 1: c.T53G, p.V18G	<i>possibly damaging</i>	nicht vollständig
5. <i>HOOK1</i>	NM_015888: Position chr1:60302551: Exon 7: c.A481T, p.S161C	<i>possibly damaging</i>	ja
6. <i>TSHR</i>	NM_000369: Position chr14:81609723: Exon 10: c.A1321C, p.T441P	<i>probably damaging</i>	ja
7. <i>FAM135A</i>	NM_020819: Position chr6:71187020: Exon 6: c.A398C, p.H133P	<i>probably damaging</i>	ja
8. <i>L2HGDH</i>	NM_024884: Position chr14:50750699: Exon 5: c.T593G, p.V198G	<i>probably damaging</i>	ja
9. <i>TTC24</i>	NM_001105669: Position chr1:156553620: Exon 6: c.T1184G, p.V395G	<i>probably damaging</i>	nicht vollständig
10. <i>SELRC1</i>	NM_023077: Position chr1:53158524: Exon 2: c.T122G, p.V41G	<i>benign</i>	nicht vollständig
11. <i>PRSS3</i>	NM_001197098: Position chr9:33798574: Exon 4: c.G524A, p.S175N	<i>benign</i>	nein
12. <i>RASSF1</i>	NM_007182: Position chr3:50375431: Exon 2: c.A262C, p.T88P	<i>unknown</i>	ja
13. <i>C17orf76-AS1</i>	NR_027166: Position chr17:16345234: Exon 5: c.A841C, p.K281Q	<i>noncoding RNA</i>	ja
14. <i>MGC45800</i>	NR_027107: Position chr4:183062660: Exon 4: c.2451_0del	<i>noncoding RNA</i>	nicht vollständig
15. <i>MIR647</i>	NR_030377: Position chr20:62574016: Exon 1: c.G64A	<i>noncoding RNA</i>	nicht vollständig
16. <i>IFFO1</i>	NR_036467: Position chr12:6649062: Exon 10: c.T2430G	<i>noncoding RNA</i>	nicht vollständig
17. <i>RPL23AP53</i>	NR_003572: Position chr8:163304: Exon 4: c.A829G	<i>noncoding RNA</i>	nicht vollständig
18. <i>LOC643387</i>	NR_026923: Position chr2:239142932: Exon 1: c.A2606G	<i>noncoding RNA</i>	nicht vollständig
19. <i>FLJ10038</i>	NR_026891: Position chr15:50646258: Exon 1: c.818_0del	<i>noncoding RNA</i>	nicht vollständig
20. <i>NECAP1</i>	NR_024260: Position chr12:8249389: Exon 7: c.T1527G	<i>noncoding RNA</i>	nicht bekannt

#### 4.6.2. Ko-Segregation der *HOOK1*-Variante und Expression

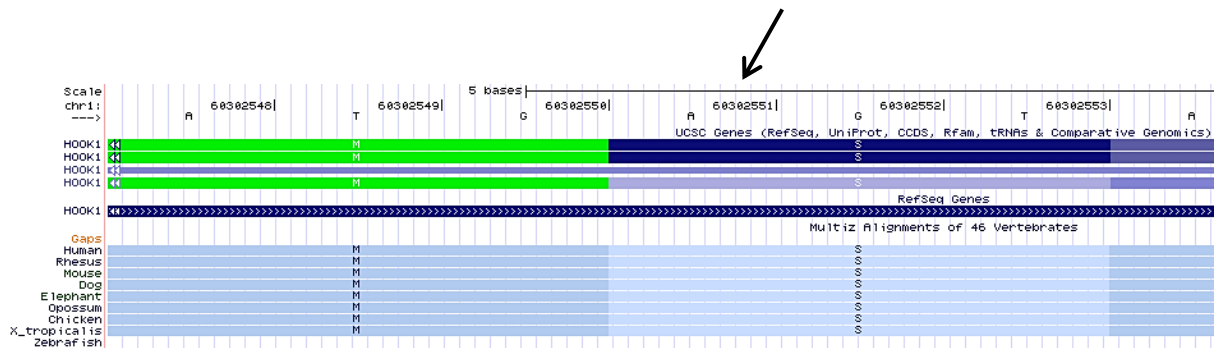
Exemplarisch wurde die Variante im Gen *HOOK1* (5.) der Tabelle 37 mittels Sanger-Sequenzierung überprüft und die Ko-Segregation in der Familie betrachtet. Alle betroffenen Familienmitglieder waren Träger der *HOOK1*-Variante. Außerdem wurde diese Variante in drei nicht betroffenen Familienmitgliedern (ID 2453, ID 2477, ID 2420) identifiziert. Trotz fehlender vollständiger Penetranz der Variante, zeigte sich eine gute Ko-Segregation (Tabelle 38).

**Tabelle 38: Ko-Segregation der *HOOK1*-Variante in der HCM-Familie 757**

ID	Phänotyp *	<i>HOOK1</i> -Variante
2417	HCM, 31 Jahre	positiv
2496	HCM, 38 Jahre	positiv
1768	HCM, 37 Jahre	positiv
2453	Nicht betroffen, 30 Jahre	positiv
1467	HCM, 71 Jahre	positiv
1834	Nicht betroffen, 32 Jahre	negativ
2418	HCM, 28 Jahre	positiv
2447	Nicht betroffen, 34 Jahre	positiv
2497	HCM, 6 Jahre	positiv
2661	Nicht betroffen, 28 Jahre	negativ
1767	HCM, 12 Jahre	positiv
1769	HCM, 15 Jahre	positiv
1833	Nicht betroffen, 11 Jahre	negativ
2419	HCM, 18 Jahre	positiv
2420	Nicht betroffen, 21 Jahre	positiv
2454	Nicht betroffen, 12 Jahre	negativ
	Betroffene mit Mutation/ Gesamtzahl Betroffene	9/9
	Nicht-Betroffene mit Mutation/ Gesamtzahl Nicht-Betroffene	3/7

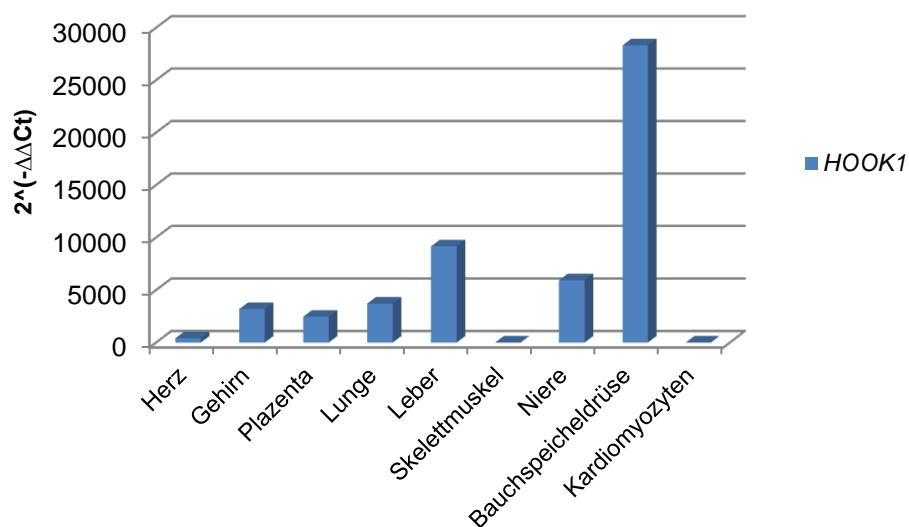
\* Bei Betroffenen Diagnosealter, bei Unbetroffenen Einschussalter

Die Variante im Gen *HOOK1* wurde durch eine PolyPhen-2 Vorhersage als möglich schädlich für das Protein eingestuft (Tabelle 37). Die Konservierung der Mutationsstelle war bis zu *Xenopus tropicalis* gegeben (Abbildung 33).



**Abbildung 33:** Konservierung der Mutationsstelle der *HOOK1*-Variante. Der Pfeil markiert die Position der Mutationsstelle im Gen *HOOK1*. Die Konservierung ist für folgende Organismen gezeigt: Mensch (*human*), Rhesusaffe (*rhesus*), Maus (*mouse*), Hund (*dog*), Elefant (*elephant*), Opossum (*opossum*), Huhn (*chicken*) und Frosch (*Xenopus tropicalis*) (<http://genome.ucsc.edu/>).

Die Expression von *HOOK1* wurde in verschiedenen Geweben/Zellen untersucht (vgl. 3.7.3 und 3.7.4). In Abbildung 34 ist die relative Expression von *HOOK1* zu sehen. Eine sehr hohe Expression wurde in der Bauchspeicheldrüse, sowie Leber und Niere beobachtet. In Herz und Kardiomyozyten konnte die Expression nicht determiniert werden (Abbildung 34).



**Abbildung 34:** Relative Expression von *HOOK1*. Als *housekeeper* wurde *GAPDH* verwendet, die Expression im Skelettmuskel diente als Kalibrator. Auf der Abszisse sind verschiedene Gewebe/Zellen aufgetragen, auf der Ordinate der 2<sup>-ΔΔCt</sup>-Wert.

#### 4.7. Gesamtauswertung der DCM- und HCM-Familien

Bei der Auswertung der einzelnen Familien fiel auf, dass identische Varianten in unterschiedlichen Familien gefunden wurden. Insgesamt 44 Varianten wurden in mindestens zwei Familien identifiziert. Dabei wurden 34 Varianten von der HCM-Familie 757 von mindestens einer DCM-Familie geteilt (Tabelle 39). Diese 44 Varianten wurden in den einzelnen Familien nicht weiter betrachtet.

**Tabelle 39: Varianten, die in mehr Familien detektiert wurden**

Gen	Annotation GRCh37/hg19	Familien
1. <i>NBPF16</i>	NM_001102663: Position: chr1:148756448: Exon 16: c.G1777A, p.G593S	009-1229, 9800, 757
2. <i>LOC388692</i>	NR_027002: Position: chr1:149281938: Exon 1: c.2463_2464insTAG	009-1229, 9800, 757
3. <i>LOC388692</i>	NR_027002: Position: chr1:149282114: Exon 1: c.T2639C	9800, 16725
4. <i>LOC388692</i>	NR_027002: Position: chr1:149287049: Exon 1: c.7574_7575insA	009-1229, 9800
5. <i>CCT3</i>	NR_036565: Position chr1:156308040: Exon 1: c.166_0del	757, 009-1229
6. <i>AK2</i>	NR_037591: Position: chr1:33475750: Exon 8: c.T1580C	009-1229, 9800, 757
7. <i>ANKRD30BL</i>	NR_027020: Position: chr2:133015184: Exon 1: c.C359G	9800, 009-1229
8. <i>ANKRD36BP2</i>	NR_015424: Position: chr2:89104898: Exon 16: c.A1433C	009-1229, 757
9. <i>GGT8P</i>	NR_003503: Position: chr2:91968909: Exon 2: c.C1240T	9800, 009-1229
10. <i>MUC4</i>	NM_018406: Position: chr3:195507226: Exon 2: c.G11225C, p.S3742T	9800, 009-1229
11. <i>FRG1</i>	NM_004477: Position: chr4:190874234: Exon 4: c.C271T, p.P91S	9800, 009-1229
12. <i>FRG1</i>	NM_004477: Position: chr4:190881933: Exon 7: c.A568G, p.K190E	009-1229, 757
13. <i>BRD8</i>	NM_006696: Position chr5:137503739: Exon 10: c.C890T, p.S297F	757, 009-1229
14. <i>LYRM2</i>	NR_028493: Position: chr6:90344779: Exon 3: c.A2678T	009-1229, 757
15. <i>TSG1</i>	NR_015362: Position: chr6:94486062: Exon 10: c.A4661C	757, 9800, 009-1229
16. <i>CFTR</i>	NM_000492: Position: chr7:117188818: Exon 10: c.1333_1336del, p.445_446del	009-1229, 757
17. <i>PRSS1</i>	NM_002769: Position: chr7:142460335: Exon 4: c.A508G, p.K170E	757, 9800
18. <i>PRSS1</i>	NM_002769: Position: chr7:142460369: Exon 4: c.G542A, p.S181N	757, 9800
19. <i>FDPSL2A</i>	NR_003262: Position: chr7:76100224: Exon 1: c.A4094G	009-1229, 757
20. <i>PABPC1</i>	NM_002568: Position: chr8:101721727: Exon 8: c.C1205T, p.P402L	009-1229, 9800, 757
21. <i>PRSS3</i>	NM_002771: Position: chr9:33797861: Exon 3: c.A235G, p.K79E	757, 9800
22. <i>ANKRD20A4</i>	NM_001098805: Position: chr9:69391171: Exon 5: c.G679A, p.A227T	009-1229, 757
23. <i>RPL23AP64</i>	NR_003040: Position: chr11:118873852: Exon 1: c.C452A	757, 009-1229
24. <i>OR4C45</i>	NM_001005513: Position: chr11:48367052: Exon 2: c.767_768insCT, p.P256fs	757, 9800

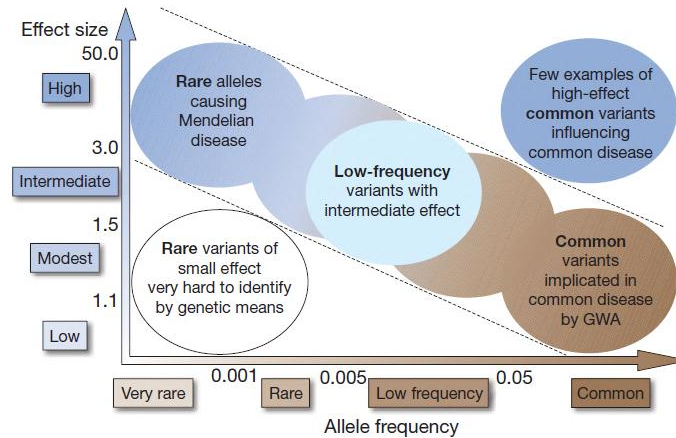
<b>Gen</b>	<b>Annotation GRCh37/hg19</b>	<b>Familien</b>
25. <i>ANKRD20A9P</i>	NR_027995: Position: chr13:19409503: Exon 1: c.A36607G	009-1229, 9800, 757
26. <i>ANKRD20A9P</i>	NR_027995: Position: chr13:19413073: Exon 1: c.T33037G	009-1229, 757
27. <i>ANKRD20A9P</i>	NR_027995: Position: chr13:19431213: Exon 1: c.C14897T	009-1229, 757
28. <i>AMDHD2</i>	NM_015944: Position: chr16:2579545: Exon 10: c.C1301G, p.A434G	9800, 009-1229
29. <i>NCOR1</i>	NM_001190440: Position: chr17:16068396: Exon 4: c.C515T, p.S172L	757, 9800
30. <i>TMEM132E</i>	NM_207313: Position: chr17:32956086: Exon 5: c.A931C, p.T311P	9800, 009-1229
31. <i>KRTAP4-7</i>	NM_033061: Position: chr17:39240627: Exon 1: c.T169C, p.S57P	757, 9800, 009-1229
32. <i>TBC1D3P2</i>	NR_027486: Position: chr17:60342111: Exon 14: c.A2018C	009-1229, 9800, 757
33. <i>TBC1D3P2</i>	NR_027486: Position: chr17:60342363: Exon 14: c.G1766C	009-1229, 757
34. <i>POTEC</i>	NM_001137671: Position: chr18:14542979: Exon 1: c.T167C, p.M56T	757, 9800
35. <i>MUC16</i>	NM_024690: Position: chr19:9002519: Exon 51: c.C40297G, p.H13433D	009-1229, 757
36. <i>FAM182B</i>	NR_026714: Position: chr20:25753792: Exon 4: c.A1923G	009-1229, 757
37. <i>FAM182A</i>	NR_026713: Position: chr20:26064080: Exon 5: c.A1638C	9800, 009-1229
38. <i>FAM182A</i>	NR_026713: Position: chr20:26064241: Exon 5: c.T1799C	757, 9800
39. <i>TP53TG5</i>	NM_014477: Position: chr20:44004159: Exon 4: c.286_288del, p.96_96del	757, 9800
40. <i>MCM8</i>	NM_032485: Position: chr20:5948171: Exon 9: c.G965A, p.S322N	757, 009-1229
41. <i>POTED</i>	NM_174981: Position: chr21:14982716: Exon 1: c.T167C, p.M56T	757, 9800
42. <i>GGT3P</i>	NR_003267: Position: chr22:18766254: Exon 8: c.G1203A	009-1229, 9800
43. <i>ABCD1</i>	NM_000033: Position: chrX:153008476: Exon 8: c.T1816C, p.S606P	009-1229, 9800, 757
44. <i>FAM104B</i>	NM_001166704: Position: chrX:55185563: Exon 2: c.A119T, p.E40V	757, 9800



## 5. Diskussion

### 5.1. Fehlende Heritabilität kardiovaskulärer Erkrankungen

Eine der zentralen Herausforderungen der modernen Genetik ist die Entschlüsselung der genetischen Ursache von Krankheiten. Monogene Erkrankungen, die einem Mendel-Erbgang folgen, werden gewöhnlich durch Mutationen mit geringen Allelfrequenzen  $<1\%$  in der Allgemeinbevölkerung verursacht (Pritchard & Cox 2002). Eine komplexe Erkrankung ist durch das Zusammenspiel von mehreren genetischen Faktoren und nicht genetischen Faktoren, wie Umwelteinflüssen, bedingt (Marian 2012). Die genetische Komponente oder allelische Architektur bezeichnet die Anzahl, Art, Häufigkeit und Effektgröße von möglich kausalen Varianten (Manolio *et al.*, 2009). Dies zeigt sich beispielsweise bei den komplexen Erkrankungen der KHK/HI, DCM und HCM. Zur Erforschung der genetischen Komponente von KHK/HI wurden in den letzten Jahren GWAS durchgeführt. Dadurch wurden die Limitationen der Kandidatengen-Studien, die hauptsächlich auf kleine Probandenzahlen und falsche Hypothesen zurückzuführen sind, umgangen. Trotz Identifizierung von 34 Suzeptibilitäts-Loci, konnten diese Varianten mit einer häufigen Allelfrequenz  $>5\%$  die Heritabilität nur in einem geringen Umfang erklären (Peden & Farrall 2011). Durch das Vorhandensein von häufigen Varianten in den Genen *PCSK9* und *LDLR* mit schwachem Effekt auf den Lipidstoffwechsel und gleichzeitigen Vorhandensein von seltenen Varianten in denjenigen Genen, die einen starken Effekt haben und die monogene Hyperlipidämie verursachen, wurde offensichtlich, dass Gene in denen häufige Varianten identifiziert wurden, auch seltene Varianten beherbergen können. Die seltenen Varianten mit Allelfrequenzen  $<0,5\%$  könnten die fehlende Heritabilität von KHK/HI erklären (Manolio *et al.*, 2009; Peden & Farrall 2011). Abbildung 35 zeigt den Zusammenhang zwischen Allelfrequenz und Effektgröße von Varianten, die durch genetische Studien identifiziert werden können. Seltene Allele mit starkem Effekt werden bei monogenen Krankheiten identifiziert, wohingegen häufige Allele mit schwachem oder mäßigem Effekt durch GWAS identifiziert werden können.



**Abbildung 35: Allelfrequenz und Effektgröße (*odds ratio*) identifizierter Varianten. Identifizierung von Varianten im Bereich der diagonalen Linien stehen im Mittelpunkt des Interesses (Manolio *et al.*, 2009).**

Bei der DCM wurde die zugrunde liegende Genetik ebenfalls durch Kandidatengenstudien, Kopplungsanalysen und kürzlich durch GWAS untersucht. In mehr als 30 Genen wurden kausale Mutationen oder Suszeptibilitäts-Loci mit einer Allelfrequenz von 1% - 8% identifiziert. Dennoch wurde die genetische Ursache nur in 30% - 35% aller familiären DCM-Fälle entschlüsselt. Wie auch bei KHK/MI wird vermutet, dass seltene Varianten mit einer Allelfrequenz  $<0,5\%$  die fehlende Heritabilität erklären können (Hershberger & Siegfried 2011). Bei HCM wurden kausale Mutationen in den Genen *MYH7* und *MYBPC3* identifiziert, die bis zu dreiviertel aller HCM-Fälle erklären. Die Identifizierung von mehr als 900 Mutationen in 23 Genen zeigen aber auch die allelische Heterogenität dieser Erkrankung, die größtenteils auf private und damit seltene Mutationen in einzelnen Familien zurückzuführen ist (Frey *et al.*, 2012).

Durch die Erkenntnisse der letzten Jahre ist eine Hypothese zur Aufklärung fehlender Heritabilität komplexer Phänotypen im Fokus. Die sog. *common disease – rare variant* Hypothese besagt, dass seltene Varianten mit starkem Effekt häufige Erkrankungen erklären. Dieser Ansatz findet nun Anwendung in der Erforschung der Genetik von komplexen Phänotypen (Marian 2012).

## 5.2. Exom-Sequenzierung zur Identifizierung seltener Varianten

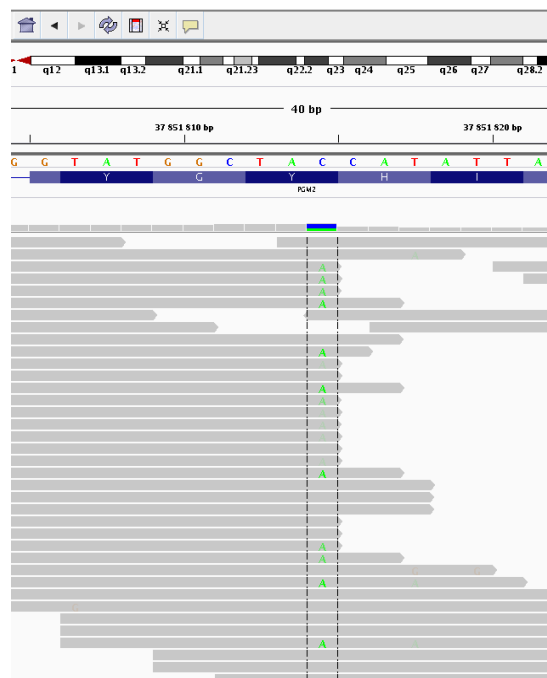
In dieser Arbeit wurde die Hypothese verfolgt, dass seltene Varianten mit einer Allelfrequenz  $<0,5\%$  und starkem Effekt die Genetik von KHK/MI, DCM und HCM weiter aufklären können. Um diese seltenen Varianten zu identifizieren, wurden Familien mit gehäuft auftretenden Krankheitsfällen ausgewählt, die ein dominantes Vererbungsmuster aufwiesen. Es wurde vermutet, dass die betroffenen Familienmitglieder gegenüber den nicht betroffenen Familienmitgliedern eine für die Familie private aber kausale Variante tragen, die somit in der Allgemeinbevölkerung nur selten oder gar nicht auftritt. Diese möglich kausalen Varianten wurden durch Exom-Sequenzierung betroffener Familienmitglieder gesucht. Exom-Sequenzierungen führten in den letzten zwei Jahren zur erfolgreichen Identifizierung kausaler Mutationen in seltenen Krankheiten und wurde so die Methode der Wahl um die Genetik von Krankheiten zu entschlüsseln (Ku *et al.*, 2012). Erst kürzlich kam es zu zwei Veröffentlichungen, in denen kausale Gene für DCM mittels Exom-Sequenzierung erfolgreich identifiziert wurden (Theis *et al.*, 2011; Norton *et al.*, 2011).

Für die Exom-Sequenzierungen in der Arbeit wurden zwei verschiedene Anreicherungsprinzipien der Exom-Bibliotheken nach Agilent Technologies und Illumina verwendet. Während bei Agilent 38 Mb Sequenzen durch Sonden abgedeckt wurden, wurden bei Illumina 62 Mb abgedeckt. Diese Diskrepanz wurde dadurch erklärt, dass die Sonden von Illumina zusätzlich 28 Mb UTR-Sequenzen umfassen (Clark *et al.*, 2011). Alle durch Sonden abgedeckten Sequenzen können aber nicht vollständig abgedeckt werden, was eine Limitation der Methode darstellt. Dies ist durch das Vorhandensein von Pseudogenen, einem niedrigen GC-Gehalt der Sequenzen und dem Auftreten von größeren Insertionen/Deletionen sowie auf eine verringerte Hybridisierungs-Effizienz zurückzuführen (Harakalova *et al.*, 2011). Voraussetzung für optimal ablaufende Exom-Sequenzierungen sind DNA-Proben mit guter Qualität. Die Anwesenheit von z.B. größeren Mengen RNA in den Proben, kann die Effizienz der DNA-Fragmentierung beeinflussen (Harakalova *et al.*, 2011). Diese Feststellung wurde bei Fragmentierung von DNA-Proben mit nicht optimaler DNA-Qualität, wie bei ID 28118 der DCM-Familie 16725 deutlich. Die Fragmentierung lieferte nicht die zu erwartenden Fragmentgrößen, wodurch eine geringe Menge an Sequenzierungsdaten erzeugt wurde. Neben der DNA-Qualität ist die Anzahl der

Amplifikationsschritte während der Herstellung der Bibliotheken entscheidend für optimale Ergebnisse. Zu viele Amplifikationsschritte können die Komplexität der Proben erhöhen, indem manche Allele verloren gehen (falsch Negative) und andere Allele überamplifiziert werden (falsch positiv Homozygote). Außerdem steigt die Anzahl von Sequenzfehlern (falsch Positive) durch mehr Amplifikationsschritte im Protokoll (Harakalova *et al.*, 2011). Um die Amplifikationsschritte bei der Herstellung der Bibliotheken nach Illumina zu reduzieren, wurde neben dem Standardprotokoll ein modifiziertes Protokoll verwendet. Hierbei wurde die Größenselektion der Fragmente durch ein Agarosegel, bei der quantitativ DNA verloren geht, ausgelassen. Dadurch konnte die Zyklenzahl der folgenden Amplifikation reduziert werden. Generell wurde bei den Proben nach modifiziertem Protokoll weniger Sequenz-Information, d.h. Menge an zugeordneter Sequenz und *coverage*, als bei den Proben des Standardprotokolls generiert. Beide Protokolle lieferten eine Überschneidung der identifizierten Varianten in den Familien. Drei der insgesamt sechs Varianten der DCM-Familie 9800 und 104 der insgesamt 201 Varianten der DCM-Familie 009-1229 wurden durch Probenvorbereitung mit beiden Protokollen identifiziert. Es konnte kein Unterschied in der Anzahl falsch positiv Homozygoter oder falsch positiver Varianten festgestellt werden. Probleme mit dem modifizierten Protokoll traten auf, wenn DNA-Proben mit nicht optimalen Fragmentgrößen vorbereitet wurden. In diesem Fall ist die Größenselektion der Fragmente ein wichtiger Schritt, um optimale Ergebnisse zu erhalten. Deshalb konnten die modifizierten Proben von ID 27974, ID 28118 und ID 28119 nicht ausgewertet werden. Die Größenselektion der Fragmente wurde bei dem Hersteller Agilent Technologies durch die Aufreinigung mit den SPRI-Kügelchen gewährleistet (Agilent Technologies, persönliche Mitteilung).

Bei der HI-Großfamilie 2001 wurden die Exom-Bibliotheken nach Agilent Technologies hergestellt, während die Exom-Bibliotheken bei den DCM- und HCM-Familien nach Illumina präpariert wurden. Tendenziell wurden bei den Sequenzierungen der Bibliotheken nach Illumina weniger Sequenzierungsdaten generiert. Im Durchschnitt wurden hier unter optimalen Bedingungen 4 Gb – 6 Gb Sequenz der Referenzsequenz zugeordnet. Dabei erreichten 65% - 77% aller Basen eine *coverage* von  $\geq 20x$ . Bei der HI-Großfamilie 2001 mit den Exom-Bibliotheken nach Agilent Technologies wurden pro Exom 6 Gb Sequenz generiert und zugeordnet, wobei 85% aller Basen eine *coverage* von  $\geq 20x$  erreichten. Die

korrekte Zuordnung der Sequenzen an die Referenzsequenz ist ein wichtiger Schritt, da Fehler zu falsch positiven Befunden führen (Day-Williams & Zeggini 2011). Die Exom-Sequenzen wurden je nach Auswertungsstrategie mit BWA oder CASAVA an das menschliche Genom *aligned*. Nach der Zuordnung der Sequenzen wurden die Varianten detektiert. Dabei können *Alignment*-Artefakte um Insertionen/Deletionen, PCR-Artefakte während der Erstellung der Exom-Bibliotheken und Fehlerprofile der *reads* zu falsch positiven Varianten führen. Durch Visualisierung des Sequenz-*Alignments* (IGV, vgl. 3.6.2) können falsch positive Varianten erkannt werden. Zum Beispiel handelt es sich oft um falsch positive Varianten, wenn sich diese am Ende von *reads* befinden, da die Fehlerrate in der Sequenz gegen Ende der *reads* steigt. Abbildung 36 zeigt die Detektion einer Variante im Gen *PGM2* in der HCM-Familie 757, wobei sich der Basenaustausch häufig gegen Ende der *reads* befindet. Diese Variante wurde durch Sanger-Sequenzierung als falsch positive Variante identifiziert.



**Abbildung 36: *Alignment*-Qualität einer Variante im Gen *PGM2*. Das Auftreten der Variante gegen Ende der *reads* ist ein Hinweis darauf, dass es sich um eine falsch positive Variante handelt.**

Die Problematik mit der steigenden Fehlerrate bei zunehmender Sequenzlänge kann umgangen werden, wenn nur eine Teil-Sequenz der *reads* für das *mapping* verwendet wird (Day-Williams & Zeggini 2011). Das wurde bei Auswertungsstrategie 2 verwirklicht, indem nur 75 bp der 100 bp langen *reads* der Referenzsequenz zugeordnet wurden. Die Detektion der Varianten wurde für beide

Auswertungsstrategien mit SAMtools durchgeführt. Für eine korrekte Detektion der Varianten ist eine hohe *coverage* für jedes Nukleotid wünschenswert (Marian 2012).

Zur Identifizierung kausaler Varianten wurden bekannte Polymorphismen, die in dbSNP132 annotiert sind, ausgeschlossen. Seltene Varianten, die als kausal in DCM-Fällen identifiziert wurden, sind mit einer niedrigen Allelfrequenz  $\leq 0,04\%$  in öffentlichen Datenbanken vertreten (Norton *et al.*, 2012). Diese Studie zeigte, dass auch bereits bekannte, seltene Varianten kausal für eine Erkrankung sein können. Bei der Auswertungsstrategie 2 wurden auch seltene Varianten mit MAF  $< 0,5\%$  der dbSNP132 ausgeschlossen, wodurch die Wahrscheinlichkeit besteht, möglich kausale Varianten übersehen zu haben. Bei der Filterung der Varianten wurden nicht-synonyme Varianten, Spleiß-Varianten und Insertionen/Deletionen im kodierenden Bereich ausgewählt. Kodierende Regionen beherbergen mehr als 85% der Mutationen von monogenen Erkrankungen (Ku *et al.*, 2012). Dennoch darf nicht vergessen werden, dass Mutationen in UTRs und nicht-kodierenden Regionen ebenfalls einen Einfluss auf Krankheiten haben können. Für autosomal-dominant vererbte Taubheit wurden z.B. Mutationen in der nicht-kodierenden miRNA, miR-96, als kausal beschrieben (Chen *et al.*, 2010). Bei der Auswertungsstrategie 2 wurden erstmals Varianten in nicht-kodierenden Regionen beibehalten, aber im Nachhinein als nicht relevant betrachtet. Mutationen, die zur Änderung der Aminosäuresequenz (nicht-synonyme Varianten) oder zur Veränderung des Leserahmens (Insertionen/Deletionen) führen, sind oft kausale Mutationen in monogenen Erkrankungen mit starkem Erscheinungsbild (Peltonen *et al.*, 2006). Deshalb wurde diese Art von Varianten bevorzugt ausgewählt. Zusätzlich wurden mit dem bioinformatischen Vorhersage-Programm PolyPhen-2 die Varianten als schädlich oder nicht schädlich für die entstehenden Proteine eingestuft. Synonyme Varianten, die in keiner Änderung der Aminosäuresequenz resultieren (stille Varianten), wurden ausgeschlossen. An dieser Stelle sei angemerkt, dass auch stille Varianten eine pathogenetische Relevanz haben können, da beispielsweise ein synonymer Polymorphismus im Gen *MDR1* (*multidrug resistance 1 gene*) die Struktur und Funktion des entstehenden Proteins beeinflusst (Kimchi-Sarfaty *et al.*, 2007).

Da Paralogie oft zu falsch positiven Befunden führt, wurden zusätzlich solche Varianten ausgeschlossen, die in paralogen Genregionen (homologe Sequenzen im gleichen Organismus) lokalisiert waren. Varianten in den Vertretern der Gen-Familie

*Mucin* wurden unter anderem häufig als paralog eingestuft. Diese Gene enthalten komplexe Insertionen/Deletionen und eine variable Anzahl an *tandem repeats* (Vinall *et al.*, 1998). Varianten, die in mehreren Familien identifiziert wurden, lagen hauptsächlich in paralogen Genen. Die korrekte Detektion von komplexen Varianten bedarf eines Algorithmus, der die Komplexität dieser Varianten berücksichtigt. Die Entwickler eines solchen Algorithmus, *Splitread*, empfehlen die Auswertung bereits gelaufener Exom-Sequenzen zu wiederholen, um komplexe Varianten mit größerer Sicherheit detektieren zu können (Karakoc *et al.*, 2012). Daher müssen auch Varianten, die mit guten Qualitätswerten ausgewählt wurden, durch Sanger-Sequenzierung validiert werden, bevor die Kausalität dieser Varianten weiter untersucht werden kann (Marian 2012).

Der Erfolg von Exom-Sequenzierungen zur Identifizierung kausaler Varianten in Familien hängt unter anderem von der Heritabilität und genetischer Heterogenität der Erkrankung ab. Exom-Sequenzierungen von entfernt-verwandt betroffenen Familienmitgliedern reduzieren die Anzahl an geteilten Varianten, wodurch die Wahrscheinlichkeit steigt, eine kausale Variante zu identifizieren (Marian 2012). Bei den Exom-Sequenzierungen der HI-Großfamilie 2001 führte diese Strategie zu einer kleinen Anzahl ( $n=4$ ) möglicher kausaler Varianten. Die Sequenzierung der nah-verwandt Betroffenen der DCM-Familie 009-1229 ergab  $n=201$  geteilte Varianten und zeigte dadurch die Wichtigkeit der Strategie. Da die Ergebnisse einer Familie aufgrund der genetischen Heterogenität nicht in anderen Familien repliziert werden können, sollte die Familie gut phänotypisiert sein und eine große Anzahl an Betroffenen aufweisen. Ist dies der Fall, kann durch Genotyp–Phänotyp Ko-Segregation eine Variante als kausal betrachtet oder ausgeschlossen werden. In Familien mit dominantem Vererbungsmuster einer Erkrankung sollten für eine vollständige Ko-Segregation alle betroffenen Familienmitglieder die Variante tragen und alle gesunden Individuen die Variante nicht aufweisen. Bei der DCM-Familie 009-1229 können nur wenige Varianten ausgeschlossen werden, da lediglich fünf Familienmitglieder diese kleine Familie repräsentieren. Zudem ist die Person der ID 04 mit zwölf Jahren sehr jung und es kann sich der Gesundheitsstatus durch eine altersabhängige Penetranz noch verändern. Dieses Beispiel zeigt, dass eine kleine Familie für den Ansatz der Exom-Sequenzierung zur Identifizierung kausaler Varianten nicht geeignet ist. Ein Problem von Ko-Segregationsanalysen stellt auch bei großen Familien die nicht vollständige Penetranz von Varianten dar, also solche

Individuen, die die Variante tragen, aber phänotypisch gesund sind (Hershberger & Siegfried 2011). Wie bereits bei den Kopplungsanalysen bei Herzinfarktfamilien erwähnt, stellen Phänokopien ebenso ein Problem dar (Schunkert *et al.*, 2010). Eine vollständige oder zumindest gute Ko-Segregation einer Variante untermauert deren Kausalität.

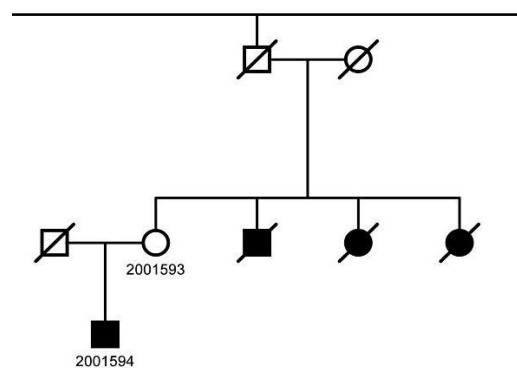
Weitere Informationen zur biologischen Plausibilität sowie bioinformatische Analysen zur funktionellen Relevanz von Varianten (PolyPhen-2) stellen weitere Hinweise auf die Kausalität dar, können diese aber nicht bestätigen (Marian 2012). Die Konservierung der Mutationsstelle wird ebenso herangezogen, um Varianten zu kategorisieren. Bei Varianten mit einem hohen Konservierungsgrad der Mutationsstelle bis hin zu niedrigen Organismen, kann es sich mit einer größeren Wahrscheinlichkeit um kausale Varianten handeln (Hershberger & Siegfried 2011). Die Kausalität der Gene, die die Varianten tragen, kann ebenso durch die Bestimmung der relativen Expression im Zielgewebe untersucht werden. Dieser Ansatz wurde bei einigen der in dieser Arbeit identifizierten Gene der DCM- und HCM-Familien durchgeführt. Eine klare Expression der Gene in Herzgewebe und Kardiomyozyten, die das Zielgewebe bzw. die Zielzellen darstellen, sollte gegeben sein. Die Expressionsdaten können somit einen Hinweis geben, ob es sich bei den Genen um mögliche Kandidatengene handelt. Letztendlich sollten aber funktionelle Studien, sowohl *in-vitro* als auch *in-vivo*, die biologische Kausalität erklären (Hershberger & Siegfried 2011; Marian 2012). Die Ergebnisse der Exom-Sequenzierung werden im Folgenden für jede Familie einzeln diskutiert.

### **5.3. Exom-Sequenzierung in einer HI-Großfamilie**

In früheren Arbeiten wurde die Genetik der HI-Großfamilie 2001 bereits mit Kopplungsanalysen untersucht, wobei kein signifikantes Kopplungssignal erhalten werden konnte. Deshalb wurde der Ansatz der Exom-Sequenzierung ausgewählt, um die genetisch kausalen Ursachen zu identifizieren. Das Exom wurde von drei betroffenen Cousins sequenziert und etwa 600 Varianten pro Individuum ausgewählt. Insgesamt wurden vier Varianten von den drei Betroffenen geteilt, die somit kausal für die Entstehung von KHK/HI sein konnten. Die Genregionen dieser Varianten wurden in den Familienmitgliedern überprüft. Keine Variante ko-segregierte vollständig mit dem Phänotyp. Wie beschrieben können fehlende



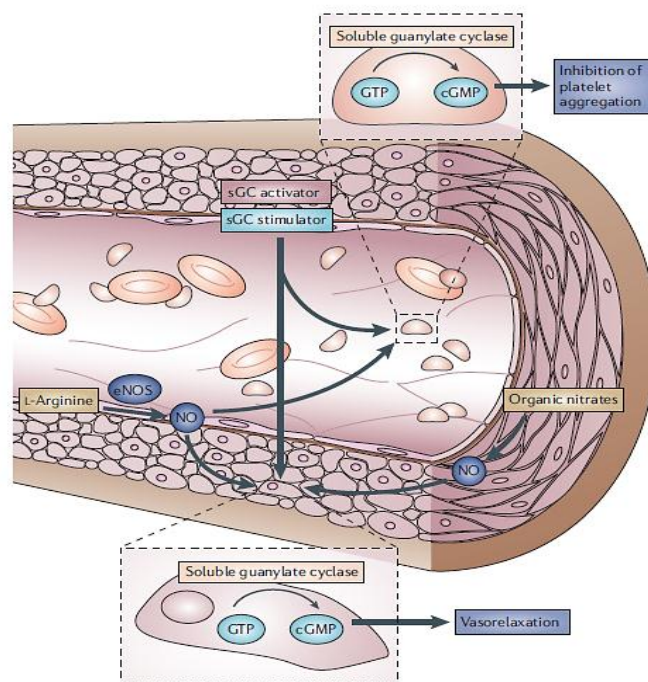
vollständige Penetranz von Varianten und Phänokopien ein Problem darstellen. In der Familie 2001 kann der Betroffene mit ID 2001594, der im Alter von 33 Jahren einen Herzinfarkt erlitten hat, als Phänokopie beschrieben werden, wenn eine vollständige Penetranz einer Variante angenommen wird (Abbildung 37). Seine Mutter mit ID 2001593 ist im Alter von 87 Jahren phänotypisch gesund, wodurch kein autosomal-dominantes Vererbungsmuster der Krankheit vorliegen kann.



**Abbildung 37: Ausschnitt aus dem Stammbaum der HI-Großfamilie 2001. Quadrat: Mann, Kreis: Frau, ausgefüllte Symbole: betroffen, offene Symbole: nicht betroffen, durchgestrichene Symbole: verstorben.**

Trotz fehlender vollständiger Ko-Segregation der Varianten in der HI-Großfamilie 2001, wurde ein Gen, *GUCY1A3*, aufgrund der in der CARDIoGRAM-Studie gezeigten Assoziation mit KHK priorisiert (Preuss *et al.*, 2010; Schunkert *et al.*, 2011). Durch die Exom-Sequenzierung wurde in diesem Gen die Insertion eines Desoxythymidinphosphat im Exon 7 identifiziert, die in einer Veränderung des Leserahmens für das entstehende Protein resultierte. Die lösliche Guanylatzyklase (sGC) besteht aus zwei Untereinheiten, die von den Genen *GUCY1A3* und *GUCY1B3* kodiert werden. Die sGC sorgt für die Stickstoffmonoxid (NO) abhängige Umwandlung von GTP zu dem sekundären Botenstoff (*second messenger*) zyklisches GMP (cGMP) (Zabel *et al.*, 1998). In mehreren Arbeiten wurde dieser Zusammenhang zwischen NO und cGMP in vaskulären glatten Muskelzellen untersucht (Ahlner *et al.*, 1991; Schmidt *et al.*, 1993). Der Anstieg an cGMP führt zu einer Abnahme des Kalzium-Spiegels in der Zelle, wodurch es zu einer reduzierten Phosphorylierung der leichten Kette von Myosin (*myosin light chain*, MLC) durch die Kalzium-abhängige Proteinkinase (Pk-C) und zu einer Abnahme der Aktivierung der MLC-Kinase (*myosin light chain kinase*, MLCK) durch Kalzium-Calmodulin-Komplexe

kommt. Dieser Mechanismus, wie auch die Aktivierung der cGMP-abhängigen Proteinkinase G und Folgereaktionen, führt zu einer Relaxation der vaskulären glatten Muskelzellen (Ahlner *et al.*, 1991; Schmidt *et al.*, 1993). Außerdem wurde eine Funktion des Signalweges in Blut-Plättchen (Thrombozyten) nachgewiesen. Die Aktivierung der cGMP-abhängigen Proteinkinase G führt zur Aktivierung von dem VASP-Protein (*vasodilator-stimulated phosphoprotein*) und dadurch zu einer Inhibition der Thrombozyten-Adhäsion und -Aggregation im Gefäßendothel (Schmidt *et al.*, 1993). Abbildung 38 zeigt den NO-sGC-cGMP Signalweg in einem Blutgefäß (Evgenov *et al.*, 2006). NO wird durch die endotheliale Stickstoffoxid-Synthase (eNOS) aus L-Arginin erzeugt. Nach Einwanderung in das Gefäßlumen und in die Gefäßwand wird sGC aktiviert. Dadurch kommt es über die beschriebenen cGMP-abhängigen Signalwege zur Vasodilatation und zur Inhibition der Thrombozyten-Aggregation (Abbildung 38).



**Abbildung 38: NO-sGC-cGMP Signalweg in einem Blutgefäß. NO wird durch eNOS erzeugt und wandert in das Gefäßlumen und in die Gefäßwand ein. Nach Aktivierung der sGC kommt es zur Vasodilatation und zur Inhibition der Thrombozyten-Aggregation (Evgenov *et al.*, 2006).**

Dieser Signalweg wurde in Zusammenhang mit der Entstehung von kardiovaskulären Krankheiten beschrieben. Aus Tierexperimenten wurde die Erkenntnis gewonnen, dass sGC ein neues pharmakologisches Ziel für die Behandlung von Atherosklerose sein könnte (Evgenov *et al.*, 2006). Neben diesen Befunden wurde erst kürzlich die

Assoziation der löslichen Guanylatzyklase mit Bluthochdruck gezeigt (Ehret *et al.*, 2011). Bluthochdruck kann wiederum zu kardiovaskulären Erkrankungen führen. Nach Definition des Hypertonus als Blutdruck  $\geq 140/90$  mm Hg oder antihypertensive Behandlung, wurden alle Mitglieder der Familie 2001 mit Bluthochdruck diagnostiziert (vgl. Tabelle 44). Da auch die nicht betroffenen Familienmitglieder an Bluthochdruck leiden, erklärt dieser Phänotyp nicht das Auftreten von KHK/HI in der Familie. Einschränkend sind diese Daten durch die teilweise nur anamnestische Erhebung nicht vollständig abgesichert. Die Literaturrecherche lieferte den Hinweis, dass es sich bei *GUCY1A3* um ein plausibles Kandidatengen für KHK/HI handelt. Dieser vermutete kausale Zusammenhang wurde auf genetischer Ebene durch das IDAHO-Screening von *GUCY1A3* in 564 HI-Patienten und 562 gesunden Individuen überprüft. Es zeigte sich, dass HI-Patienten mit einer 5-fach höheren Wahrscheinlichkeit eine durch PolyPhen-2 vorhergesagte, schädliche Variante im Gen tragen als gesunde Individuen.

Trotz der vorhandenen Assoziation von *GUCY1A3* mit KHK, dem plausiblen Signalweg und einer gehäuften Mutationsrate von *GUCY1A3* in HI-Patienten konnte diese Variante die Entstehung von KHK/HI in der Familie nicht vollständig erklären. Deshalb wurde die Hypothese aufgestellt, dass zwei oder mehr Gene über Interaktion der durch die Varianten geschädigten Proteine die Erkrankung KHK/HI in der Familie 2001 verursachen. Die Funktion der durch Exom-Sequenzierung identifizierten Varianten-tragenden Gene *ETFDH*, *GCLC* und *CCT7*, wurde durch Literaturrecherche beleuchtet. Es wurde beschrieben, dass das Gen *ETFDH* für die in der inneren Mitochondrien-Membran lokalisierte *electron-transferring-flavoprotein dehydrogenase* ETFDH kodiert, welche als Elektronen-Carrier für das *electron-transfer flavoprotein* ETF fungiert und Ubiquinon reduziert. Nachgewiesen wurde, dass beide Proteine in der  $\beta$ -Oxidation von CoA-Fettsäure-Derivaten beteiligt sind (Ruzicka & Beinert 1977). Außerdem wurde ein Zusammenhang von Mutationen in den Genen *ETF/ETFDH* mit der Krankheit Multiple Acyl-CoA Dehydrogenase-Defekte (Glutarazidurie Typ II) gezeigt (Olsen *et al.*, 2003). Mutationen im Gen *GCLC*, kodierend für die katalytische Untereinheit der Glutamat-Cystein-Ligase GCL, zeigten in früheren Arbeiten eine Assoziation mit hämolytischer Anämie und Herzinfarkt (Meister & Anderson 1983; Koide *et al.*, 2003; Willis *et al.*, 2011). Das Protein, das für das Gen *CCT7* (*chaperonin containing TCP1, subunit 7 (eta)*) kodiert, wurde als Chaperon beschrieben, das mit weiteren Proteinen den Chaperonin-

enthaltenden-TCP1-Komplex bildet. Dieser Komplex zeigt eine ATP-abhängige Faltung von Proteinen, wie Aktin und Tubulin (Kubota *et al.*, 1994). Ein interessantes Ergebnis lieferte die Arbeit von Hanafy *et al.*, in der die Interaktion von CCT7 mit der Guanylatzyklase gezeigt wurde. CCT7 bindet an sGC und sorgt für eine verringerte NO-Aktivierung der Guanylatzyklase (Hanafy *et al.*, 2004). Um die Hypothese interagierender Kandidatengene auf genetischer Ebene zu testen, wurden die kodierenden Exone von *ETFDH*, *GCLC* und *CCT7* in 48 HI-Patienten sequenziert. Hierbei wurden keine weiteren Varianten identifiziert. Es besteht durchaus die Möglichkeit, dass die Anzahl der sequenzierten Individuen zu klein war, um weitere genetische Variationen zu entdecken.

In der HI-Großfamilie 2001 zeigte die Variante im Gen *CCT7* die beste Ko-Segregation. 11 der 15 Betroffenen waren Träger der *CCT7*-Variante. Dennoch wurden durch Sequenzierung zusätzlicher HI-Patienten keine weiteren Varianten in dem Gen identifiziert, um die Kausalität von *CCT7* für KHK/HI zu untermauern. Die Variante im Gen *GUCY1A3* wurde nur von 8 der 15 betroffenen Familienmitglieder geteilt. Das Gen konnte aber durch andere Anhaltspunkte als plausibles Kandidatengen angesehen werden. Die Interaktion von *CCT7* und sGC könnte nun die fehlende Kausalität jeder einzelnen Variante erklären. In der Familie waren von 7 Trägern der *CCT7*-Einzelvariante, 3 Individuen gesund und 4 Individuen erkrankt. Die 2 Träger der *GUCY1A3*-Einzelvariante waren gesund. Interessanterweise waren alle 7 Doppelmutationsträger der *CCT7*- und *GUCY1A3*-Variante erkrankt. Dadurch zeigte sich, dass beide Varianten zusammen immer zur Ausprägung der Krankheit führen. PD. Dr. Marcus Fischer, der in früheren Arbeiten die Kopplungsanalyse in der Familie bereits untersucht hat, hat diese mit einem Zwei-Locus-Modell erneut berechnet. Durch persönliche Kommunikation wurde mitgeteilt, dass beide Loci einen maximalen LOD-score von 5,68 ergeben. Jeder einzelne Locus zeigte einen nicht signifikanten LOD-score (LOD-score von 1,11 für *CCT7* und 0,08 für *GUCY1A3*), was bereits durch die gerechnete Kopplungsanalyse bekannt war. Der Zwei-Locus-Kopplungsbefund untermauert die gemeinsame Bedeutung von *CCT7* und *GUCY1A3*. Somit scheint der NO-sGC-cGMP Signalweg auf genetischer Ebene einen Einfluss auf die Entstehung von KHK/HI zu haben. Diese Vermutung muss nun mit funktionellen Studien belegt werden.

#### 5.4. Exom-Sequenzierung in DCM- und HCM-Familien

Wie bei der HI-Großfamilie wurde die Exom-Sequenzierung auch in den DCM- und HCM-Familien genutzt, um genetisch kausale Varianten zu identifizieren. Fünf betroffene Individuen der DCM-Familie 9800 wurden für die Exom-Sequenzierung ausgewählt. Es handelte sich um zwei Cousinen und einen Cousin und um eine Nichte und einen Neffen von einer Cousine. Es wurden fünf nicht-kodierende Varianten und eine kodierende Variante im Gen *PASD1* identifiziert. Das Gen *PASD1* wurde einem Tumor-Hoden-Antigen (*cancer/testis antigen*) zugeordnet, das sich auf Krebszellen befindet (Sahota *et al.*, 2006). Nach Betrachtung der Funktion des Gens, stellte es kein plausibles Kandidatengen für DCM mehr dar. Da Phänokopien auch in dieser Familie nicht ausgeschlossen werden können, besteht die Möglichkeit, die Auswertung mit vier der fünf sequenzierten Exome durchzuführen. Bei der ID 19208 kann es sich um eine Phänokopie handeln, wenn eine vollständige Penetranz einer Variante angenommen wird, da die Mutter nicht betroffen ist und somit kein autosomal-dominantes Vererbungsmuster vorliegen kann. Falls diese Auswertung dennoch zu keinem interessanten Kandidatengen führt, kann vermutet werden, dass die kausale Variante in einer Region liegt, die nicht oder mit niedriger *coverage* abgedeckt wurde und dadurch nicht identifiziert werden konnte.

In der DCM-Familie 16725 wurden vier entfernt-verwandte Familienmitglieder für die Exom-Sequenzierung ausgewählt. Zusätzlich wurde ein Nicht-Familienmitglied der ID 29134 als Kontroll-Individuum sequenziert. Die Auswertung ergab acht von den Familienmitgliedern geteilte Varianten, die nicht in ID 29134 identifiziert wurden. Drei Varianten in nicht-kodierenden RNAs wurden aufgrund des vorliegenden Phänotyps als nicht relevant angesehen. Die Variante im Gen *MUC4* wurde ebenfalls nicht weiter betrachtet, da dieses Gen einen hohen Anteil an paralogen Sequenzen aufweist. Zwei Varianten in *MUC4* wurden vorangehend in der Familie als paralog eingestuft. Konservierte Varianten mit starkem Effekt auf die entstehenden Proteine können den Phänotyp am wahrscheinlichsten erklären. Deshalb wurden zwei identifizierte Varianten, eine Deletion von drei Basenpaaren im Gen *KCNJ12* und eine *stopgain*-Variante im Gen *TTN* priorisiert. In der Literatur wurden Mutationen im Titin-Gen bereits als kausal für DCM beschrieben (Gerull *et al.*, 2002; Gerull *et al.*, 2006). Nach Betrachtung der Ko-Segregation der *TTN-stopgain*-Variante in der Familie zeigte sich eine mangelnde Penetranz der Variante. 7 der 11 nicht

betroffenen Familienmitglieder wiesen die Variante auf. Eine *stopgain*-Variante im Gen wurde bereits im Jahr 2006 in einer Familie mit DCM identifiziert, die eine gute Ko-Segregation zeigte: von den 13 nicht betroffenen Familienmitgliedern wurde die Variante nur in einer Person identifiziert (Gerull *et al.*, 2006). Aufgrund der auch in anderen Arbeiten beschriebener starker Penetranz von Titin-Mutationen, konnte die hier beschriebene Variante nicht vollständig die Entstehung von DCM in der Familie 16725 erklären. Im Allgemeinen können *modifier*-Gene die fehlende Penetranz von Mutationen erklären. Ein *modifier*-Gen verändert durch eine Variante im Gen den Schweregrad und den Fortschritt einer Krankheit unabhängig von der primären Ursache (Mestroni & Taylor 2011). Zum Beispiel wurde durch Sequenzierung des Promotors von Calmodulin III (*CALM3*) ein neuer Polymorphismus als *modifier* für HCM identifiziert (Friedrich *et al.*, 2009). Bei Calmodulin III handelt es sich um ein plausibles Kandidatengen, da Proteine, die die Kalzium-Homöostase regulieren, eine wichtige Rolle bei der Entstehung von HCM spielen (Chiu *et al.*, 2007). Neben *modifier*-Effekten könnte auch die zweite Variante in der Familie, die Variante im Gen *KCNJ12*, die fehlende Penetranz/Genetik erklären. Es wurde beschrieben, dass *KCNJ12* für einen Kalium-Kanal kodiert, der die Ionen ins Zellinnere einschleust (Wible *et al.*, 1995). Dieser Kanal trägt zum kardialen nach innen gerichteten Ionenstrom bei (*cardiac inward rectifier current, IK1*). Es handelt sich somit um ein plausibles Kandidatengen, weshalb zunächst die Ko-Segregation der Variante in der Familie betrachtet werden sollte. Mit funktionellen Studien sollte anschließend eine mögliche Kausalität der *TTN*-Variante sowie der *KCNJ12*-Variante überprüft werden.

Von zwei betroffenen Familienmitgliedern, der Großmutter und ihrem Enkel, der DCM-Familie 009-1229 wurde das Exom sequenziert. Da es sich um eine kleine Familie handelte, konnten keine entfernt-verwandten Personen, wie z.B. Cousin/Cousine rekrutiert werden. Aufgrund des nahen Verwandtschaftsverhältnisses der sequenzierten Personen wurden 201 geteilte Varianten identifiziert, von denen 129 in kodierenden Regionen lagen. Bislang wurde exemplarisch die Ko-Segregation von neun Varianten überprüft. Nur die Variante im Gen *WHAMM* zeigte eine vollständige Ko-Segregation. Alle betroffenen Familienmitglieder waren gegenüber den nicht betroffenen Familienmitgliedern Träger der Variante. Es wurde beschrieben, dass das Gen *WHAMM* (*WAS protein homolog associated with actin, golgi membranes and microtubule*) für ein NPF (*nucleation-promoting factor*) kodiert, der die Polymerisation von Aktin stimuliert. Zudem besitzt das Protein eine Domäne

für Golgi-Membran-Assoziation und eine Domäne für die Bindung von Mikrotubuli und spielt dadurch eine wichtige Rolle bei der Dynamik des Zytoskeletts (Campellone *et al.*, 2008). Laut einer PolyPhen-2 Vorhersage wurde die Variante im Gen *WHAMM* als nicht schädlich für das entstehende Protein eingestuft. Zudem war die Konservierung der Mutationsstelle nicht vollständig gegeben. Dadurch wird die Kausalität dieser Variante, trotz vorhandener Ko-Segregation in der Familie und biologischer Plausibilität in Frage gestellt. Die Ko-Segregation in der Familie zu betrachten, gestaltet sich wie bereits erwähnt durch ID 04 als schwierig. Mit Hilfe von Expressionsdaten der Gene kann die Kausalität von Varianten neben der Ko-Segregation untersucht werden. Eine vorhandene Expression der Gene in Herzgewebe und Kardiomyozyten sollte gegeben sein. Exemplarisch wurden die Expressionsdaten für drei Gene, *GJB4*, *ZG16* und *INTS1* gezeigt. Von den Varianten dieser Gene wurde ein starker Effekt auf die entstehenden Proteine vermutet. *GJB4* und *ZG16* zeigten keine oder nur schwache Expression im Zielgewebe (Herzgewebe und Kardiomyozyten) und können deshalb als Kandidatengene ausgeschlossen werden. Das Gen *INTS1* (*integrator complex subunit 1*) wurde ubiquitär exprimiert. Dieses Gen kodiert für eine Untereinheit des sog. Integrator-Komplexes, der mit der C-terminalen Domäne der RNA-Polymerase II assoziiert (Baillat *et al.*, 2005). Durch die biologische Funktion des Gens erscheint die ubiquitäre Expression plausibel, liefert aber somit keinen Hinweis auf Kausalität. Die Exom-Sequenzierung von entfernt-verwandten Familienmitgliedern konnte, wie bei der DCM-Familie 16725, zur Identifizierung von wenigen Kandidatengenen führen. Dies konnte bei der DCM-Familie 009-1229 nicht erreicht werden.

Neben den drei DCM-Familien wurde eine Familie mit gehäuft auftretender HCM für die Exom-Sequenzierung ausgewählt. Von zwei betroffenen Cousins und einem Großcousin der Familie wurde das Exom sequenziert. Insgesamt wurden nach Filterung 20 Varianten von den drei betroffenen Familienmitgliedern geteilt. Dabei handelte es sich um zwei Deletionen und zehn SNVs im kodierenden Bereich und um acht Varianten in nicht-kodierenden Regionen. Bislang wurde die Ko-Segregation der Variante im Gen *HOOK1*, die durch eine PolyPhen-2 Vorhersage als möglich schädlich für das Protein eingestuft wurde, in der Familie untersucht. Alle neun betroffenen Familienmitglieder und drei der sieben nicht betroffenen Familienmitglieder waren Träger der *HOOK1*-Variante. Trotz fehlender vollständiger Penetranz der Variante, zeigte sich eine gute Ko-Segregation. Es wurde gezeigt,

dass HOOK1 ein Vertreter der Mikrotubulus-bindenden Proteine ist und eine Rolle beim Transport von internalisierten Proteinen zum Endosom spielt (Weimer *et al.*, 2005). Da eine biologische Plausibilität vorliegen könnte, wurde die Expression des Gens in verschiedenen Geweben/Zellen betrachtet. Im Zielgewebe (Herzgewebe und Kardiomyozyten) konnte keine Expression determiniert werden. Dadurch wurde das Gen als nicht plausibles Kandidatengen eingestuft. Für die weitere Auswertung sollte die Ko-Segregation der restlichen Varianten in kodierenden Regionen überprüft werden. Die Variante im Gen *RYR3* würde aufgrund der Funktion des Gens ein plausibles Kandidatengen darstellen. Das Gen wurde als kodierend für den Ryanodin-Rezeptor RyR3 beschrieben, der als Kalziumkanal fungiert und in vielen Geweben, einschließlich Herz und glatten Muskelzellen exprimiert wird (Jiang *et al.*, 2003). Die Gesamtauswertung der Ergebnisse von den einzelnen Familien zeigte, dass 34 Varianten in der HCM-Familie und in einer oder zwei DCM-Familien gefunden wurden. Aus der Literatur ist bekannt, dass Mutationen in den gleichen Genen sowohl DCM als auch HCM verursachen können (Harvey & Leinwand 2011). Die Position und Art der Mutation im Gen beeinflusst dabei den kardialen Phänotyp. Neue Genotyp-Phänotyp Korrelationen wurden zum Beispiel zu Mutationen in *MYH7* und *MYBPC3* untersucht (Waldmüller *et al.*, 2011). Deshalb sollten diese Varianten nicht ausgeschlossen werden und weiterhin als möglich kausale Varianten in Betracht gezogen werden.

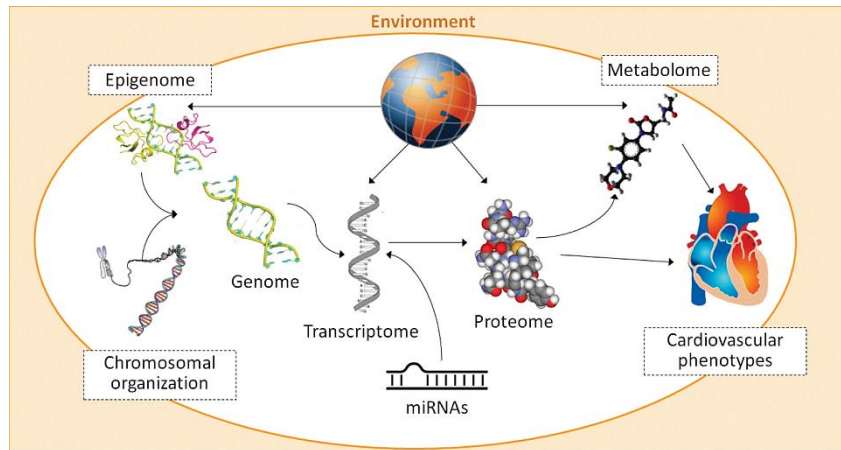
Die Exom-Sequenzierung in DCM- und HCM-Familien führte ebenfalls zur Identifizierung von möglichen Kandidatengenen. Varianten in den Genen *TTN* und *KCNJ12* wurden als möglich kausal für DCM und die Variante im Gen *RYR3* als möglich kausal für HCM identifiziert. Die biologische Plausibilität dieser Gene ist ausser Frage gestellt, so dass im nächsten Schritt die Kausalität der Varianten überprüft werden sollte.



## 5.5. Ausblick

Die Funktion von 98,5% des Genoms, mit 3,2 Milliarden Nukleotiden, ist bislang ungeklärt. Ungefähr 1,5% des Genoms sind mit einer starken Beeinflussung von biologischen und klinischen Phänotypen impliziert (Marian 2012). Durch die Entwicklung der Hochdurchsatz-Sequenzierung wurden neue Wege eröffnet, die fehlende Genetik von klinischen Phänotypen zu untersuchen. Neben der Exom-Sequenzierung kann auch die Genom-Sequenzierung durch fallende Sequenzierungskosten schon bald in größerem Umfang durchgeführt werden. Dadurch wird die personalisierte Genomforschung ermöglicht (Ansorge 2009). Bereits im Jahr 2010 wurde das Genom eines Menschen im klinischen Kontext analysiert (Ashley *et al.*, 2010). Der untersuchte Patient wies eine positive Familienanamnese für kardiovaskuläre Erkrankungen auf. Die Analyse der Varianten ergab ein erhöhtes genetisches Risiko für Herzinfarkt, Diabetes und für manche Arten von Krebs. Die Arbeit zeigte, dass durch die Genom-Sequenzierung klinisch relevante Informationen für Einzelpatienten erhalten werden können. Dennoch ist die personalisierte Medizin auf dieser Ebene noch weit entfernt vom diagnostischen Routine-Einsatz. Herausforderungen der Genom-Sequenzierung gehen mit Problemen, wie der akkuraten Zuordnung der Sequenzen und korrekten Detektion und Auswahl von Varianten einher. Es gilt diese zu überwinden, bevor das „persönliche Genom“ in die klinische Praxis eintritt (Samani *et al.*, 2010).

Durch die neuen Sequenzierungsmethoden wurde, neben der Genom-Sequenzierung, die Erforschung des Epigenoms erleichtert (Schnabel *et al.*, 2012). Epigenetische Modifikationen, wie DNA-Methylierung, Histon-Modifikation und Nukleosom-Positionierung, kontrollieren die Genexpression unabhängig von genetischen Varianten. DNA-Methylierung kann z.B. durch Bisulfat-Sequenzierung untersucht werden. RNA-Sequenzierung ermöglicht die Erforschung des Transkriptoms, d.h. der qualitativen und quantitativen Unterschiede der Genexpression in Zellen. Abbildung 39 zeigt das komplexe Netzwerk aus Genom, Epigenom, Transkriptom, Proteom und Metabolom, das erforscht werden muss, um die Genetik von z.B. kardiovaskulären Phänotypen besser zu verstehen.



**Abbildung 39: Komplexität kardiovaskulärer Phänotypen. Das Zusammenspiel von Genom, Epigenom, Transkriptom, Proteom und Metabolom hat eine Bedeutung bei der Entstehung von kardiovaskulären Phänotypen (Schnabel *et al.*, 2012).**

Durch die neuen Sequenzierungsmethoden werden bessere Einblicke in die Komplexität dieses Netzwerkes gewonnen. Dadurch wird die Erforschung der Genetik kardiovaskulärer Erkrankungen revolutioniert.

## 6. Zusammenfassung

In dieser Arbeit wurde die Genetik von kardiovaskulären Krankheiten untersucht. Dafür wurden Familien mit gehäuft auftretenden Krankheitsfällen von koronarer Herzerkrankung/Herzinfarkt, dilatativer Kardiomyopathie oder hypertrophischer Kardiomyopathie rekrutiert. Die Exom-Sequenzierung von betroffenen Familienmitgliedern sollte zur Identifizierung der genetischen Ursache in den Familien führen. Seltene Varianten, die von Familienmitgliedern geteilt wurden, wird ein starker Effekt zugeschrieben und daher als möglich kausal betrachtet.

Bei der HI-Großfamilie wurden vier mögliche Varianten identifiziert, von denen zwei nach Betrachtung der Kausalität als plausibel angesehen wurden. Es handelte sich um Varianten in den Genen *GUCY1A3*, das für die  $\alpha$ -Untereinheit der löslichen Guanylatzyklase kodiert, und *CCT7*, das für ein Chaperon des Chaperonin-enthaltenden-TCP1-Komplexes kodiert. Durch die Identifizierung kausaler Varianten in den Genen wurde der NO-sGC-cGMP Signalweg als neu für die Entstehung von KHK/HI vorgeschlagen.

Die Exom-Sequenzierung von Mitgliedern einer DCM-Familie führte zur Identifizierung einer neuen Variante im Titin-Gen, die in der Einführung eines Stoppkodons resultierte. Aufgrund schwacher Penetranz dieser Variante in der Familie wird aber vermutet, dass eine zweite Variante (*modifier*) zur Entstehung der DCM in dieser Familie beiträgt. Die Variante im Gen *KCNJ12*, das für einen Kaliumkanal kodiert, wird dabei aufgrund der biologischen Plausibilität besonders ins Auge gefasst. Bei der untersuchten HCM-Familie wurden ebenso potentiell kausale Varianten identifiziert, wobei hier eine Variante im Gen *RYR3*, das für einen Ryanodin-Rezeptor kodiert, aufgrund der biologischen Plausibilität priorisiert wird. Die Exom-Sequenzierung in zwei weiteren DCM-Familien führte bislang nicht zur Identifizierung kausaler Varianten. Dieses Ergebnis spiegelt die Komplexität und die Limitationen der Methode wider.

Die Exom-Sequenzierung wurde bisher erfolgreich als Methode genutzt, um kausale Varianten in monogenen Erkrankungen zu identifizieren. Diese Arbeit zeigt, dass die Exom-Sequenzierung auch in komplexen Erkrankungen zur Identifizierung kausaler Varianten führen kann.

## 7. Literaturverzeichnis

- Abbate R., Sticchi E., Fatini C. Genetics of cardiovascular disease. *Clinical Cases in Mineral and Bone Metabolism* **5** (1), 63-66 (2008).
- Ahlner J., Andersson R. G., Torfgard K., Axelsson K. L. Organic nitrate esters: clinical use and mechanisms of actions. *Pharmacol.Rev.* **43**, 351-423 (1991).
- Andresdottir M. B., Sigurdsson G., Sigvaldason H., Gudnason V. Fifteen percent of myocardial infarctions and coronary revascularizations explained by family history unrelated to conventional risk factors. The Reykjavik Cohort Study. *Eur.Heart J.* **23**, 1655-1663 (2002).
- Ansorge W., EMBL Heidelberg Process for sequencing nucleic acids without gel sieving media on solid support and DNA chips (Verfahren zur Sequenzierung von Nukleinsäuren ohne Gele). *German Patent Application and Corresponding Worldwide Patent Applications* **DE 41 41 178 A1**, (1991).
- Ansorge W. J. Next-generation DNA sequencing techniques. *N.Biotechnol.* **25**, 195-203 (2009).
- Ashley E. A., Butte A. J., Wheeler M. T., Chen R., Klein T. E., Dewey F. E., Dudley J. T., Ormond K. E., Pavlovic A., Morgan A. A., Pushkarev D., Neff N. F., Hudgins L., Gong L., Hodges L. M., Berlin D. S., Thorn C. F., Sangkuhl K., Hebert J. M., Woon M., Sagreiya H., Whaley R., Knowles J. W., Chou M. F., Thakuria J. V., Rosenbaum A. M., Zaranek A. W., Church G. M., Greely H. T., Quake S. R., Altman R. B. Clinical assessment incorporating a personal genome. *Lancet* **375**, 1525-1535 (2010).
- Baillat D., Hakimi M. A., Naar A. M., Shilatifard A., Cooch N., Shiekhattar R. Integrator, a multiprotein mediator of small nuclear RNA processing, associates with the C-terminal repeat of RNA polymerase II. *Cell* **123**, 265-276 (2005).
- Broeckel U., Hengstenberg C., Mayer B., Holmer S., Martin L. J., Comuzzie A. G., Blangero J., Nurnberg P., Reis A., Riegger G. A., Jacob H. J., Schunkert H. A comprehensive linkage analysis for myocardial infarction and its related risk factors. *Nat.Genet.* **30**, 210-214 (2002).
- Cambien F., Poirier O., Lecerf L., Evans A., Cambou J. P., Arveiler D., Luc G., Bard J. M., Bara L., Ricard S., . Deletion polymorphism in the gene for angiotensin-converting enzyme is a potent risk factor for myocardial infarction. *Nature* **359**, 641-644 (1992).
- Campellone K. G., Webb N. J., Znameroski E. A., Welch M. D. WHAMM is an Arp2/3 complex activator that binds microtubules and functions in ER to Golgi transport. *Cell* **134**, 148-161 (2008).

- Chen J. M., Ferec C., Cooper D. N. Revealing the human mutome. *Clin.Genet.* **78**, 310-320 (2010).
- Chiu C., Tebo M., Ingles J., Yeates L., Arthur J. W., Lind J. M., Semsarian C. Genetic screening of calcium regulation genes in familial hypertrophic cardiomyopathy. *J.Mol.Cell Cardiol.* **43**, 337-343 (2007).
- Clark M. J., Chen R., Lam H. Y., Karczewski K. J., Chen R., Euskirchen G., Butte A. J., Snyder M. Performance comparison of exome DNA sequencing technologies. *Nat.Biotechnol.* **29**, 908-914 (2011).
- Danecek P., Auton A., Abecasis G., Albers C. A., Banks E., DePristo M. A., Handsaker R. E., Lunter G., Marth G. T., Sherry S. T., McVean G., Durbin R. The variant call format and VCFtools. *Bioinformatics.* **27**, 2156-2158 (2011).
- Day-Williams A. G., Zeggini E. The effect of next-generation sequencing technology on complex trait research. *Eur.J.Clin.Invest* **41**, 561-567 (2011).
- Doggen C. J., Kunz G., Rosendaal F. R., Lane D. A., Vos H. L., Stubbs P. J., Manger C., V, Ireland H. A mutation in the thrombomodulin gene, 127G to A coding for Ala25Thr, and the risk of myocardial infarction in men. *Thromb.Haemost.* **80**, 743-748 (1998).
- Duboscq-Bidot L., Charron P., Ruppert V., Fauchier L., Richter A., Tavazzi L., Arbustini E., Wichter T., Maisch B., Komajda M., Isnard R., Villard E. Mutations in the ANKRD1 gene encoding CARP are responsible for human dilated cardiomyopathy. *Eur.Heart J.* **30**, 2128-2136 (2009).
- Dudbridge F., Gusnanto A. Estimation of significance thresholds for genomewide association scans. *Genet.Epidemiol.* **32**, 227-234 (2008).
- Ehret G. B., Munroe P. B., Rice K. M., Bochud M., Johnson A. D., Chasman D. I., Smith A. V., Tobin M. D., Verwoert G. C., Hwang S. J., Pihur V., Vollenweider P., O'Reilly P. F., Amin N., Bragg-Gresham J. L., Teumer A., Glazer N. L., Launer L., Zhao J. H., Aulchenko Y., Heath S., Sober S., Parsa A., Luan J., Arora P., Dehghan A., Zhang F., Lucas G., Hicks A. A., Jackson A. U., Peden J. F., Tanaka T., Wild S. H., Rudan I., Igl W., Milaneschi Y., Parker A. N., Fava C., Chambers J. C., Fox E. R., Kumari M., Go M. J., van der Harst P., Kao W. H., Sjogren M., Vinay D. G., Alexander M., Tabara Y., Shaw-Hawkins S., Whincup P. H., Liu Y., Shi G., Kuusisto J., Tayo B., Seielstad M., Sim X., Nguyen K. D., Lehtimaki T., Matullo G., Wu Y., Gaunt T. R., Onland-Moret N. C., Cooper M. N., Platou C. G., Org E., Hardy R., Dahgam S., Palmen J., Vitart V., Braund P. S., Kuznetsova T., Uiterwaal C. S., Adeyemo A., Palmas W., Campbell H., Ludwig B., Tomaszewski M., Tzoulaki I., Palmer N. D., Aspelund T., Garcia M., Chang Y. P., O'Connell J. R., Steinle N. I., Grobbee D. E., Arking D. E., Kardina S. L., Morrison A. C., Hernandez D., Najjar S., McArdle W. L., Hadley D., Brown M. J., Connell J. M., Hingorani A. D., Day I. N., Lawlor D. A., Beilby J. P., Lawrence R. W., Clarke R., Hopewell J. C., Ongen H., Dreisbach A. W., Li Y., Young J. H., Bis J. C., Kahonen M., Viikari J., Adair L. S., Lee N. R., Chen M. H., Olden M., Pattaro C., Bolton J. A., Kottgen A., Bergmann S.,

- Mooser V., Chaturvedi N., Frayling T. M., Islam M., Jafar T. H., Erdmann J., Kulkarni S. R., Bornstein S. R., Grassler J., Groop L., Voight B. F., Kettunen J., Howard P., Taylor A., Guarrera S., Ricceri F., Emilsson V., Plump A., Barroso I., Khaw K. T., Weder A. B., Hunt S. C., Sun Y. V., Bergman R. N., Collins F. S., Bonnycastle L. L., Scott L. J., Stringham H. M., Peltonen L., Perola M., Vartiainen E., Brand S. M., Staessen J. A., Wang T. J., Burton P. R., Artigas M. S., Dong Y., Snieder H., Wang X., Zhu H., Lohman K. K., Rudock M. E., Heckbert S. R., Smith N. L., Wiggins K. L., Doumatey A., Shriner D., Veldre G., Viigimaa M., Kinra S., Prabhakaran D., Tripathy V., Langefeld C. D., Rosengren A., Thelle D. S., Corsi A. M., Singleton A., Forrester T., Hilton G., McKenzie C. A., Salako T., Iwai N., Kita Y., Ogiwara T., Ohkubo T., Okamura T., Ueshima H., Umemura S., Eyheramendy S., Meitinger T., Wichmann H. E., Cho Y. S., Kim H. L., Lee J. Y., Scott J., Sehmi J. S., Zhang W., Hedblad B., Nilsson P., Smith G. D., Wong A., Narisu N., Stancakova A., Raffel L. J., Yao J., Kathiresan S., O'Donnell C. J., Schwartz S. M., Ikram M. A., Longstreth W. T., Jr., Mosley T. H., Seshadri S., Shrine N. R., Wain L. V., Morken M. A., Swift A. J., Laitinen J., Prokopenko I., Zitting P., Cooper J. A., Humphries S. E., Danesh J., Rasheed A., Goel A., Hamsten A., Watkins H., Bakker S. J., van Gilst W. H., Janipalli C. S., Mani K. R., Yajnik C. S., Hofman A., Mattace-Raso F. U., Oostra B. A., Demirkan A., Isaacs A., Rivadeneira F., Lakatta E. G., Orru M., Scuteri A., Ala-Korpela M., Kangas A. J., Lyytikäinen L. P., Soininen P., Tukiainen T., Wurtz P., Ong R. T., Dorr M., Kroemer H. K., Volker U., Volzke H., Galan P., Hercberg S., Lathrop M., Zelenika D., Deloukas P., Mangino M., Spector T. D., Zhai G. Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk. *Nature* **478**, 103-109 (2011).
- Erdmann J., Grosshennig A., Braund P. S., König I. R., Hengstenberg C., Hall A. S., Linsel-Nitschke P., Kathiresan S., Wright B., Tregouet D. A., Cambien F., Bruse P., Aherrahrou Z., Wagner A. K., Stark K., Schwartz S. M., Salomaa V., Elosua R., Melander O., Voight B. F., O'Donnell C. J., Peltonen L., Siscovick D. S., Altshuler D., Merlini P. A., Peyvandi F., Bernardinelli L., Ardissino D., Schillert A., Blankenberg S., Zeller T., Wild P., Schwarz D. F., Tiret L., Perret C., Schreiber S., El Mokhtari N. E., Schafer A., Marz W., Renner W., Bugert P., Kluter H., Schrezenmeier J., Rubin D., Ball S. G., Balmforth A. J., Wichmann H. E., Meitinger T., Fischer M., Meisinger C., Baumert J., Peters A., Ouwehand W. H., Deloukas P., Thompson J. R., Ziegler A., Samani N. J., Schunkert H. New susceptibility locus for coronary artery disease on chromosome 3q22.3. *Nat. Genet.* **41**, 280-282 (2009).
- Erdmann J., Schunkert H. Genetik der koronaren Herzkrankheit und des Herzinfarkts. *Medizinische Genetik* **3**, 316-320 (2007).
- Evgenov O. V., Pacher P., Schmidt P. M., Hasko G., Schmidt H. H., Stasch J. P. NO-independent stimulators and activators of soluble guanylate cyclase: discovery and therapeutic potential. *Nat. Rev. Drug Discov.* **5**, 755-768 (2006).
- Ewing B., Hillier L., Wendl M. C., Green P. Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Res.* **8**, 175-185 (1998).
- Fischer M., Broeckel U., Holmer S., Baessler A., Hengstenberg C., Mayer B., Erdmann J., Klein G., Riegger G., Jacob H. J., Schunkert H. Distinct heritable

patterns of angiographic coronary artery disease in families with myocardial infarction. *Circulation* **111**, 855-862 (2005).

Frey N., Luedde M., Katus H. A. Mechanisms of disease: hypertrophic cardiomyopathy. *Nature Reviews Cardiology* **9**, 91-100 (2012).

Friedrich F. W., Bausero P., Sun Y., Treszl A., Kramer E., Juhr D., Richard P., Wegscheider K., Schwartz K., Brito D., Arbustini E., Waldenstrom A., Isnard R., Komajda M., Eschenhagen T., Carrier L. A new polymorphism in human calmodulin III gene promoter is a potential modifier gene for familial hypertrophic cardiomyopathy. *Eur.Heart J.* **30**, 1648-1655 (2009).

Gardemann A., Lohre J., Katz N., Tillmanns H., Hehrlein F. W., Haberbosch W. The 4G4G genotype of the plasminogen activator inhibitor 4G/5G gene polymorphism is associated with coronary atherosclerosis in patients at high risk for this disease. *Thromb.Haemost.* **82**, 1121-1126 (1999).

Geisterfer-Lowrance A. A., Kass S., Tanigawa G., Vosberg H. P., McKenna W., Seidman C. E., Seidman J. G. A molecular basis for familial hypertrophic cardiomyopathy: a beta cardiac myosin heavy chain gene missense mutation. *Cell* **62**, 999-1006 (1990).

Gerull B., Atherton J., Geupel A., Sasse-Klaassen S., Heuser A., Frenneaux M., McNabb M., Granzier H., Labeit S., Thierfelder L. Identification of a novel frameshift mutation in the giant muscle filament titin in a large Australian family with dilated cardiomyopathy. *J.Mol.Med.(Berl)* **84**, 478-483 (2006).

Gerull B., Gramlich M., Atherton J., McNabb M., Trombitas K., Sasse-Klaassen S., Seidman J. G., Seidman C., Granzier H., Labeit S., Frenneaux M., Thierfelder L. Mutations of TTN, encoding the giant muscle filament titin, cause familial dilated cardiomyopathy. *Nat.Genet.* **30**, 201-204 (2002).

Gilissen C., Arts H. H., Hoischen A., Spruijt L., Mans D. A., Arts P., van L. B., Steehouwer M., van R. J., Kant S. G., Roepman R., Knoers N. V., Veltman J. A., Brunner H. G. Exome sequencing identifies WDR35 variants involved in Sensenbrenner syndrome. *Am.J.Hum.Genet.* **87**, 418-423 (2010).

Glass C. K., Witztum J. L. Atherosclerosis. the road ahead. *Cell* **104**, 503-516 (2001).

Hanafy K. A., Martin E., Murad F. CCTeta, a novel soluble guanylyl cyclase-interacting protein. *J.Biol.Chem.* **279**, 46946-46953 (2004).

Harakalova M., Mokry M., Hrdlickova B., Renkens I., Duran K., van R. H., Lansu N., van R. M., de B. E., Nijman I. J., Kloosterman W. P., Cuppen E. Multiplexed array-based and in-solution genomic enrichment for flexible and cost-effective targeted next-generation sequencing. *Nat.Protoc.* **6**, 1870-1886 (2011).

- Harvey P. A., Leinwand L. A. The cell biology of disease: cellular mechanisms of cardiomyopathy. *J.Cell Biol.* **194**, 355-365 (2011).
- Helgadottir A., Manolescu A., Thorleifsson G., Gretarsdottir S., Jonsdottir H., Thorsteinsdottir U., Samani N. J., Gudmundsson G., Grant S. F., Thorgeirsson G., Sveinbjornsdottir S., Valdimarsson E. M., Matthiasson S. E., Johannsson H., Gudmundsdottir O., Gurney M. E., Sainz J., Thorhallsdottir M., Andresdottir M., Frigge M. L., Topol E. J., Kong A., Gudnason V., Hakonarson H., Gulcher J. R., Stefansson K. The gene encoding 5-lipoxygenase activating protein confers risk of myocardial infarction and stroke. *Nat.Genet.* **36**, 233-239 (2004).
- Helgadottir A., Thorleifsson G., Manolescu A., Gretarsdottir S., Blondal T., Jonasdottir A., Jonasdottir A., Sigurdsson A., Baker A., Palsson A., Masson G., Gudbjartsson D. F., Magnusson K. P., Andersen K., Levey A. I., Backman V. M., Matthiasdottir S., Jonsdottir T., Palsson S., Einarsdottir H., Gunnarsdottir S., Gylfason A., Vaccarino V., Hooper W. C., Reilly M. P., Granger C. B., Austin H., Rader D. J., Shah S. H., Quyyumi A. A., Gulcher J. R., Thorgeirsson G., Thorsteinsdottir U., Kong A., Stefansson K. A common variant on chromosome 9p21 affects the risk of myocardial infarction. *Science* **316**, 1491-1493 (2007).
- Hershberger R. E., Parks S. B., Kushner J. D., Li D., Ludwigsen S., Jakobs P., Nauman D., Burgess D., Partain J., Litt M. Coding sequence mutations identified in MYH7, TNNT2, SCN5A, CSRP3, LBD3, and TCAP from 313 patients with familial or idiopathic dilated cardiomyopathy. *Clin.Transl.Sci.* **1**, 21-26 (2008).
- Hershberger R. E., Siegfried J. D. Update 2011: clinical and genetic issues in familial dilated cardiomyopathy. *J.Am.Coll.Cardiol.* **57**, 1641-1649 (2011).
- Hoischen A., van Bon B. W., Gilissen C., Arts P., van L. B., Steehouwer M., de V. P., de R. R., Wieskamp N., Mortier G., Devriendt K., Amorim M. Z., Revencu N., Kidd A., Barbosa M., Turner A., Smith J., Oley C., Henderson A., Hayes I. M., Thompson E. M., Brunner H. G., de Vries B. B., Veltman J. A. De novo mutations of SETBP1 cause Schinzel-Giedion syndrome. *Nat.Genet.* **42**, 483-485 (2010).
- Horne B. D., Camp N. J., Muhlestein J. B., Cannon-Albright L. A. Identification of excess clustering of coronary heart diseases among extended pedigrees in a genealogical population database. *Am.Heart J.* **152**, 305-311 (2006).
- Jefferies J. L., Towbin J. A. Dilated cardiomyopathy. *Lancet* **375**, 752-762 (2010).
- Jiang D., Xiao B., Li X., Chen S. R. Smooth muscle tissues express a major dominant negative splice variant of the type 3 Ca<sup>2+</sup> release channel (ryanodine receptor). *J.Biol.Chem.* **278**, 4763-4769 (2003).
- Kaiser J. DNA sequencing. A plan to capture human diversity in 1000 genomes. *Science* **319**, 395 (2008).



Karakoc E., Alkan C., O'Roak B. J., Dennis M. Y., Vives L., Mark K., Rieder M. J., Nickerson D. A., Eichler E. E. Detection of structural variants and indels within exome data. *Nat.Methods* **9**, 176-178 (2012).

Kathiresan S., Voight B. F., Purcell S., Musunuru K., Ardissino D., Mannucci P. M., Anand S., Engert J. C., Samani N. J., Schunkert H., Erdmann J., Reilly M. P., Rader D. J., Morgan T., Spertus J. A., Stoll M., Girelli D., McKeown P. P., Patterson C. C., Siscovick D. S., O'Donnell C. J., Elosua R., Peltonen L., Salomaa V., Schwartz S. M., Melander O., Altshuler D., Ardissino D., Merlini P. A., Berzuini C., Bernardinelli L., Peyvandi F., Tubaro M., Celli P., Ferrario M., Fétique R., Marziliano N., Casari G., Galli M., Ribichini F., Rossi M., Bernardi F., Zoncin P., Piazza A., Mannucci P. M., Schwartz S. M., Siscovick D. S., Yee J., Friedlander Y., Elosua R., Marrugat J., Lucas G., Subirana I., Sala J., Ramos R., Kathiresan S., Meigs J. B., Williams G., Nathan D. M., MacRae C. A., O'Donnell C. J., Salomaa V., Havulinna A. S., Peltonen L., Melander O., Berglund G., Voight B. F., Kathiresan S., Hirschhorn J. N., Asselta R., Duga S., Sreafico M., Musunuru K., Daly M. J., Purcell S., Voight B. F., Purcell S., Nemesh J., Korn J. M., McCarroll S. A., Schwartz S. M., Yee J., Kathiresan S., Lucas G., Subirana I., Elosua R., Surti A., Guiducci C., Gianniny L., Mirel D., Parkin M., Burtt N., Gabriel S. B., Samani N. J., Thompson J. R., Braund P. S., Wright B. J., Balmforth A. J., Ball S. G., Hall A. S., Schunkert H., Erdmann J., Linsel-Nitschke P., Lieb W., Ziegler A., König I., Hengstenberg C., Fischer M., Stark K., Grosshennig A., Preuss M., Wichmann H. E., Schreiber S., Schunkert H., Samani N. J., Erdmann J., Ouwehand W., Hengstenberg C., Deloukas P., Scholz M., Cambien F., Reilly M. P., Li M., Chen Z., Wilensky R., Matthai W., Qasim A., Hakonarson H. H., Devaney J., Burnett M. S., Pichard A. D., Kent K. M., Satler L., Lindsay J. M., Waksman R., Knouff C. W., Waterworth D. M., Walker M. C., Mooser V., Epstein S. E., Rader D. J., Scheffold T., Berger K., Stoll M., Häge A., Girelli D., Martinelli N., Olivieri O., Corrocher R., Morgan T., Spertus J. A., McKeown P., Patterson C. C., Schunkert H., Erdmann E., Linsel-Nitschke P., Lieb W., Ziegler A., König I. R., Hengstenberg C., Fischer M., Stark K., Grosshennig A., Preuss M., Wichmann H. E., Schreiber S., Holm H., Thorleifsson G., Thorsteinsdottir U., Stefansson K., Engert J. C., Do R., Xie C., Anand S., Kathiresan S., Ardissino D., Mannucci P. M., Siscovick D., O'Donnell C. J., Samani N. J., Melander O., Elosua R., Peltonen L., Salomaa V., Schwartz S. M., Altshuler D. Genome-wide association of early-onset myocardial infarction with single nucleotide polymorphisms and copy number variants. *Nat.Genet.* **41**, 334-341 (2009).

Kimchi-Sarfaty C., Oh J. M., Kim I. W., Sauna Z. E., Calcagno A. M., Ambudkar S. V., Gottesman M. M. A "silent" polymorphism in the MDR1 gene changes substrate specificity. *Science* **315**, 525-528 (2007).

Kimura A. Molecular etiology and pathogenesis of hereditary cardiomyopathy. *Circ.J.* **72 Suppl A**, A38-A48 (2008).

Koide S., Kugiyama K., Sugiyama S., Nakamura S., Fukushima H., Honda O., Yoshimura M., Ogawa H. Association of polymorphism in glutamate-cysteine ligase catalytic subunit gene with coronary vasomotor dysfunction and myocardial infarction. *J.Am.Coll.Cardiol.* **41**, 539-545 (2003).

- Ku C. S., Cooper D. N., Polychronakos C., Naidoo N., Wu M., Soong R. Exome sequencing: dual role as a discovery and diagnostic tool. *Ann.Neurol.* **71**, 5-14 (2012).
- Kubota H., Hynes G., Carne A., Ashworth A., Willison K. Identification of six Tcpl-1-related genes encoding divergent subunits of the TCP-1-containing chaperonin. *Curr.Biol.* **4**, 89-99 (1994).
- Lander E. S., Linton L. M., Birren B., Nusbaum C., Zody M. C., Baldwin J., Devon K., Dewar K., Doyle M., FitzHugh W., Funke R., Gage D., Harris K., Heaford A., Howland J., Kann L., Lehoczky J., LeVine R., McEwan P., McKernan K., Meldrim J., Mesirov J. P., Miranda C., Morris W., Naylor J., Raymond C., Rosetti M., Santos R., Sheridan A., Sougnez C., Stange-Thomann N., Stojanovic N., Subramanian A., Wyman D., Rogers J., Sulston J., Ainscough R., Beck S., Bentley D., Burton J., Clee C., Carter N., Coulson A., Deadman R., Deloukas P., Dunham A., Dunham I., Durbin R., French L., Grafham D., Gregory S., Hubbard T., Humphray S., Hunt A., Jones M., Lloyd C., McMurray A., Matthews L., Mercer S., Milne S., Mullikin J. C., Mungall A., Plumb R., Ross M., Shownkeen R., Sims S., Waterston R. H., Wilson R. K., Hillier L. W., McPherson J. D., Marra M. A., Mardis E. R., Fulton L. A., Chinwalla A. T., Pepin K. H., Gish W. R., Chissoe S. L., Wendl M. C., Delehaunty K. D., Miner T. L., Delehaunty A., Kramer J. B., Cook L. L., Fulton R. S., Johnson D. L., Minx P. J., Clifton S. W., Hawkins T., Branscomb E., Predki P., Richardson P., Wenning S., Slezak T., Doggett N., Cheng J. F., Olsen A., Lucas S., Elkin C., Uberbacher E., Frazier M., Gibbs R. A., Muzny D. M., Scherer S. E., Bouck J. B., Sodergren E. J., Worley K. C., Rives C. M., Gorrell J. H., Metzker M. L., Naylor S. L., Kucherlapati R. S., Nelson D. L., Weinstock G. M., Sakaki Y., Fujiyama A., Hattori M., Yada T., Toyoda A., Itoh T., Kawagoe C., Watanabe H., Totoki Y., Taylor T., Weissenbach J., Heilig R., Saurin W., Artiguenave F., Brottier P., Bruls T., Pelletier E., Robert C., Wincker P., Smith D. R., Doucette-Stamm L., Rubenfield M., Weinstock K., Lee H. M., Dubois J., Rosenthal A., Platzer M., Nyakatura G., Taudien S., Rump A., Yang H., Yu J., Wang J., Huang G., Gu J., Hood L., Rowen L., Madan A., Qin S., Davis R. W., Federspiel N. A., Abola A. P., Proctor M. J., Myers R. M., Schmutz J., Dickson M., Grimwood J., Cox D. R., Olson M. V., Kaul R., Raymond C., Shimizu N., Kawasaki K., Minoshima S., Evans G. A., Athanasiou M., Schultz R., Roe B. A., Chen F., Pan H., Ramser J., Lehrach H., Reinhardt R., McCombie W. R., de la Bastide M., Dedhia N., Blocker H., Hornischer K., Nordsiek G., Agarwala R., Aravind L., Bailey J. A., Bateman A., Batzoglou S., Birney E., Bork P., Brown D. G., Burge C. B., Cerutti L., Chen H. C., Church D., Clamp M., Copley R. R., Doerks T., Eddy S. R., Eichler E. E., Furey T. S., Galagan J., Gilbert J. G., Harmon C., Hayashizaki Y., Haussler D., Hermjakob H., Hokamp K., Jang W., Johnson L. S., Jones T. A., Kasif S., Kasprzyk A., Kennedy S., Kent W. J., Kitts P., Koonin E. V., Korf I., Kulp D., Lancet D., Lowe T. M., McLysaght A., Mikkelsen T., Moran J. V., Mulder N., Pollara V. J., Ponting C. P., Schuler G., Schultz J., Slater G., Smit A. F., Stupka E., Szustakowski J., Thierry-Mieg D., Thierry-Mieg J., Wagner L., Wallis J., Wheeler R., Williams A., Wolf Y. I., Wolfe K. H., Yang S. P., Yeh R. F., Collins F., Guyer M. S., Peterson J., Felsenfeld A., Wetterstrand K. A., Patrinos A., Morgan M. J., de J. P., Catanese J. J., Osoegawa K., Shizuya H., Choi S., Chen Y. J. Initial sequencing and analysis of the human genome. *Nature* **409**, 860-921 (2001).

- Lazaruk K., Wang Y., Zhong J., Maltchenko S., Rabkin S., Hunkapiller K., Furtado M., Petrauskene O., Guegler K., Gilbert D., Spier E. The Design Process for a New Generation of Quantitative Gene Expression Analysis Tools. *Applied Biosystems* (2006).
- Li H., Handsaker B., Wysoker A., Fennell T., Ruan J., Homer N., Marth G., Abecasis G., Durbin R. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. **25**, 2078-2079 (2009).
- Lieb W., Mayer B., Konig I. R., Borwitzky I., Gotz A., Kain S., Hengstenberg C., Linsel-Nitschke P., Fischer M., Doring A., Wichmann H. E., Meitinger T., Kreutz R., Ziegler A., Schunkert H., Erdmann J. Lack of association between the MEF2A gene and myocardial infarction. *Circulation* **117**, 185-191 (2008).
- Liu Y., Niu W., Wu Z., Su X., Chen Q., Lu L., Jin W. Variants in exon 11 of MEF2A gene and coronary artery disease: evidence from a case-control study, systematic review, and meta-analysis. *PLoS.One*. **7**, e31406 (2012).
- Lupski J. R., Reid J. G., Gonzaga-Jauregui C., Rio D. D., Chen D. C., Nazareth L., Bainbridge M., Dinh H., Jing C., Wheeler D. A., McGuire A. L., Zhang F., Stankiewicz P., Halperin J. J., Yang C., Gehman C., Guo D., Irikat R. K., Tom W., Fantin N. J., Muzny D. M., Gibbs R. A. Whole-genome sequencing in a patient with Charcot-Marie-Tooth neuropathy. *N.Engl.J.Med.* **362**, 1181-1191 (2010).
- Manolio T. A., Collins F. S., Cox N. J., Goldstein D. B., Hindorff L. A., Hunter D. J., McCarthy M. I., Ramos E. M., Cardon L. R., Chakravarti A., Cho J. H., Guttmacher A. E., Kong A., Kruglyak L., Mardis E., Rotimi C. N., Slatkin M., Valle D., Whittemore A. S., Boehnke M., Clark A. G., Eichler E. E., Gibson G., Haines J. L., Mackay T. F., McCarroll S. A., Visscher P. M. Finding the missing heritability of complex diseases. *Nature* **461**, 747-753 (2009).
- Marenberg M. E., Risch N., Berkman L. F., Floderus B., de F. U. Genetic susceptibility to death from coronary heart disease in a study of twins. *N.Engl.J.Med.* **330**, 1041-1046 (1994).
- Marian A. J. Hypertrophic cardiomyopathy: from genetics to treatment. *Eur.J.Clin.Invest* **40**, 360-369 (2010).
- Marian A. J. Molecular genetic studies of complex phenotypes. *Transl.Res.* **159**, 64-79 (2012).
- McPherson R., Pertsemlidis A., Kavaslar N., Stewart A., Roberts R., Cox D. R., Hinds D. A., Pennacchio L. A., Tybjaerg-Hansen A., Folsom A. R., Boerwinkle E., Hobbs H. H., Cohen J. C. A common allele on chromosome 9 associated with coronary heart disease. *Science* **316**, 1488-1491 (2007).
- Meister A., Anderson M. E. Glutathione. *Annu.Rev.Biochem.* **52**, 711-760 (1983).

- Mestroni L., Taylor M. R. Pharmacogenomics, personalized medicine, and heart failure. *Discov.Med.* **11**, 551-561 (2011).
- Mullis K. B., Faloona F. A. Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. *Methods Enzymol.* **155**, 335-350 (1987).
- Ng S. B., Bigam A. W., Buckingham K. J., Hannibal M. C., McMillin M. J., Gildersleeve H. I., Beck A. E., Tabor H. K., Cooper G. M., Mefford H. C., Lee C., Turner E. H., Smith J. D., Rieder M. J., Yoshiura K., Matsumoto N., Ohta T., Niikawa N., Nickerson D. A., Bamshad M. J., Shendure J. Exome sequencing identifies MLL2 mutations as a cause of Kabuki syndrome. *Nat.Genet.* **42**, 790-793 (2010a).
- Ng S. B., Buckingham K. J., Lee C., Bigam A. W., Tabor H. K., Dent K. M., Huff C. D., Shannon P. T., Jabs E. W., Nickerson D. A., Shendure J., Bamshad M. J. Exome sequencing identifies the cause of a mendelian disorder. *Nat.Genet.* **42**, 30-35 (2010b).
- Ng S. B., Turner E. H., Robertson P. D., Flygare S. D., Bigam A. W., Lee C., Shaffer T., Wong M., Bhattacharjee A., Eichler E. E., Bamshad M., Nickerson D. A., Shendure J. Targeted capture and massively parallel sequencing of 12 human exomes. *Nature* **461**, 272-276 (2009).
- Norton N., Li D., Rieder M. J., Siegfried J. D., Rampersaud E., Zuchner S., Mangos S., Gonzalez-Quintana J., Wang L., McGee S., Reiser J., Martin E., Nickerson D. A., Hershberger R. E. Genome-wide studies of copy number variation and exome sequencing identify rare variants in BAG3 as a cause of dilated cardiomyopathy. *Am.J.Hum.Genet.* **88**, 273-282 (2011).
- Norton N., Robertson P. D., Rieder M. J., Zuchner S., Rampersaud E., Martin E., Li D., Nickerson D. A., Hershberger R. E. Evaluating Pathogenicity of Rare Variants from Dilated Cardiomyopathy in the Exome Era. *Circ.Cardiovasc.Genet.* (2012).
- O'Donnell C. J., Nabel E. G. Genomics of cardiovascular disease. *N.Engl.J.Med.* **365**, 2098-2109 (2011).
- Olsen R. K., Andresen B. S., Christensen E., Bross P., Skovby F., Gregersen N. Clear relationship between ETF/ETFDH genotype and phenotype in patients with multiple acyl-CoA dehydrogenation deficiency. *Hum.Mutat.* **22**, 12-23 (2003).
- Olson T. M., Michels V. V., Thibodeau S. N., Tai Y. S., Keating M. T. Actin mutations in dilated cardiomyopathy, a heritable form of heart failure. *Science* **280**, 750-752 (1998).
- Peden J. F., Farrall M. Thirty-five common variants for coronary artery disease: the fruits of much collaborative labour. *Hum.Mol.Genet.* **20**, R198-R205 (2011).

- Peltonen L., Perola M., Naukkarinen J., Palotie A. Lessons from studying monogenic disease for common disease. *Hum.Mol.Genet.* **15 Spec No 1**, R67-R74 (2006).
- Pfaffl M. W. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res.* **29**, e45 (2001).
- Preuss M., König I. R., Thompson J. R., Erdmann J., Absher D., Assimes T. L., Blankenberg S., Boerwinkle E., Chen L., Cupples L. A., Hall A. S., Halperin E., Hengstenberg C., Holm H., Laaksonen R., Li M., Marz W., McPherson R., Musunuru K., Nelson C. P., Burnett M. S., Epstein S. E., O'Donnell C. J., Quertermous T., Rader D. J., Roberts R., Schillert A., Stefansson K., Stewart A. F., Thorleifsson G., Voight B. F., Wells G. A., Ziegler A., Kathiresan S., Reilly M. P., Samani N. J., Schunkert H. Design of the Coronary ARtery Disease Genome-Wide Replication And Meta-Analysis (CARDIoGRAM) Study: A Genome-wide association meta-analysis involving more than 22 000 cases and 60 000 controls. *Circ.Cardiovasc.Genet.* **3**, 475-483 (2010).
- Pritchard J. K., Cox N. J. The allelic architecture of human disease genes: common disease-common variant...or not? *Hum.Mol.Genet.* **11**, 2417-2423 (2002).
- Rader D. J., Cohen J., Hobbs H. H. Monogenic hypercholesterolemia: new insights in pathogenesis and treatment. *J.Clin.Invest* **111**, 1795-1803 (2003).
- Robinson J. T., Thorvaldsdottir H., Winckler W., Guttman M., Lander E. S., Getz G., Mesirov J. P. Integrative genomics viewer. *Nat.Biotechnol.* **29**, 24-26 (2011).
- Robinson P. S. Detection System for PCR Assay. *United States Patent US 7,615,620 B2*, (2009).
- 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., Fullerton H. J., Gillespie C., Hailpern S. M., Heit J. A., Howard V. J., Kissela B. M., Kittner S. J., Lackland D. T., Lichtman J. H., Lisabeth L. D., Makuc D. M., Marcus G. M., Marelli A., Matchar D. B., Moy C. S., Mozaffarian D., Mussolino M. E., Nichol G., Paynter N. P., Soliman E. Z., Sorlie P. D., Sotoodehnia N., Turan T. N., Virani S. S., Wong N. D., Woo D., Turner M. B. Heart disease and stroke statistics--2012 update: a report from the american heart association. *Circulation* **125**, e2-e220 (2012).
- Ross R. Atherosclerosis--an inflammatory disease. *N.Engl.J.Med.* **340**, 115-126 (1999).
- Ruzicka F. J., Beinert H. A new iron-sulfur flavoprotein of the respiratory chain. A component of the fatty acid beta oxidation pathway. *J.Biol.Chem.* **252**, 8440-8445 (1977).
- Saffitz J. E. Dependence of electrical coupling on mechanical coupling in cardiac myocytes: insights gained from cardiomyopathies caused by defects in cell-cell connections. *Ann.N.Y.Acad.Sci.* **1047**, 336-344 (2005).

- Sahota S. S., Goonewardena C. M., Cooper C. D., Liggins A. P., Ait-Tahar K., Zojer N., Stevenson F. K., Banham A. H., Pulford K. PASD1 is a potential multiple myeloma-associated antigen. *Blood* **108**, 3953-3955 (2006).
- Saiki R. K., Scharf S., Faloona F., Mullis K. B., Horn G. T., Erlich H. A., Arnheim N. Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science* **230**, 1350-1354 (1985).
- Samani N. J., Burton P., Mangino M., Ball S. G., Balmforth A. J., Barrett J., Bishop T., Hall A. A genomewide linkage study of 1,933 families affected by premature coronary artery disease: The British Heart Foundation (BHF) Family Heart Study. *Am.J.Hum.Genet.* **77**, 1011-1020 (2005).
- Samani N. J., Erdmann J., Hall A. S., Hengstenberg C., Mangino M., Mayer B., Dixon R. J., Meitinger T., Braund P., Wichmann H. E., Barrett J. H., König I. R., Stevens S. E., Szymczak S., Tregouet D. A., Iles M. M., Pahlke F., Pollard H., Lieb W., Cambien F., Fischer M., Ouwehand W., Blankenberg S., Balmforth A. J., Baessler A., Ball S. G., Strom T. M., Braenne I., Gieger C., Deloukas P., Tobin M. D., Ziegler A., Thompson J. R., Schunkert H. Genomewide association analysis of coronary artery disease. *N.Engl.J.Med.* **357**, 443-453 (2007).
- Samani N. J., Tomaszewski M., Schunkert H. The personal genome--the future of personalised medicine? *Lancet* **375**, 1497-1498 (2010).
- Sanger F., Nicklen S., Coulson A. R. DNA sequencing with chain-terminating inhibitors. *Proc.Natl.Acad.Sci.U.S.A* **74**, 5463-5467 (1977).
- Satoh M., Takahashi M., Sakamoto T., Hiroe M., Marumo F., Kimura A. Structural analysis of the titin gene in hypertrophic cardiomyopathy: identification of a novel disease gene. *Biochem.Biophys.Res.Comm.* **262**, 411-417 (1999).
- Schmidt H. H., Lohmann S. M., Walter U. The nitric oxide and cGMP signal transduction system: regulation and mechanism of action. *Biochim.Biophys.Acta* **1178**, 153-175 (1993).
- Schnabel R. B., Baccarelli A., Lin H., Ellinor P. T., Benjamin E. J. Next steps in cardiovascular disease genomic research--sequencing, epigenetics, and transcriptomics. *Clin.Chem.* **58**, 113-126 (2012).
- Schunkert H., Erdmann J., Samani N. J. Genetics of myocardial infarction: a progress report. *Eur.Heart J.* **31**, 918-925 (2010).
- Schunkert H., König I. R., Kathiresan S., Reilly M. P., Assimes T. L., Holm H., Preuss M., Stewart A. F., Barbalic M., Gieger C., Absher D., Aherrahrou Z., Allayee H., Altshuler D., Anand S. S., Andersen K., Anderson J. L., Ardissino D., Ball S. G., Balmforth A. J., Barnes T. A., Becker D. M., Becker L. C., Berger K., Bis J. C., Boekholdt S. M., Boerwinkle E., Braund P. S., Brown M. J., Burnett M. S., Buyschaert I., Carlquist J. F., Chen L., Cichon S., Codd V., Davies R. W.,

Dedoussis G., Dehghan A., Demissie S., Devaney J. M., Diemert P., Do R., Doering A., Eifert S., Mokhtari N. E., Ellis S. G., Elosua R., Engert J. C., Epstein S. E., de F. U., Fischer M., Folsom A. R., Freyer J., Gigante B., Girelli D., Gretarsdottir S., Gudnason V., Gulcher J. R., Halperin E., Hammond N., Hazen S. L., Hofman A., Horne B. D., Illig T., Iribarren C., Jones G. T., Jukema J. W., Kaiser M. A., Kaplan L. M., Kastelein J. J., Khaw K. T., Knowles J. W., Kolovou G., Kong A., Laaksonen R., Lambrechts D., Leander K., Lettre G., Li M., Lieb W., Loley C., Lotery A. J., Mannucci P. M., Maouche S., Martinelli N., McKeown P. P., Meisinger C., Meitinger T., Melander O., Merlini P. A., Mooser V., Morgan T., Muhleisen T. W., Muhlestein J. B., Munzel T., Musunuru K., Nahrstaedt J., Nelson C. P., Nothen M. M., Olivieri O., Patel R. S., Patterson C. C., Peters A., Peyvandi F., Qu L., Quyyumi A. A., Rader D. J., Rallidis L. S., Rice C., Rosendaal F. R., Rubin D., Salomaa V., Sampietro M. L., Sandhu M. S., Schadt E., Schafer A., Schillert A., Schreiber S., Schrezenmeir J., Schwartz S. M., Siscovick D. S., Sivananthan M., Sivapalaratnam S., Smith A., Smith T. B., Snopce J. D., Soranzo N., Spertus J. A., Stark K., Stirrups K., Stoll M., Tang W. H., Tennstedt S., Thorgeirsson G., Thorleifsson G., Tomaszewski M., Uitterlinden A. G., van Rij A. M., Voight B. F., Wareham N. J., Wells G. A., Wichmann H. E., Wild P. S., Willenborg C., Witteman J. C., Wright B. J., Ye S., Zeller T., Ziegler A., Cambien F., Goodall A. H., Cupples L. A., Quertermous T., Marz W., Hengstenberg C., Blankenberg S., Ouwehand W. H., Hall A. S., Deloukas P., Thompson J. R., Stefansson K., Roberts R., Thorsteinsdottir U., O'Donnell C. J., McPherson R., Erdmann J., Samani N. J. Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease. *Nat.Genet.* **43**, 333-338 (2011).

Schuster S. C. Next-generation sequencing transforms today's biology. *Nat.Methods* **5**, 16-18 (2008).

Seidman C. E., Seidman J. G. Identifying sarcomere gene mutations in hypertrophic cardiomyopathy: a personal history. *Circ.Res.* **108**, 743-750 (2011).

Sharp P. A., Sugden B., Sambrook J. Detection of two restriction endonuclease activities in *Haemophilus parainfluenzae* using analytical agarose-ethidium bromide electrophoresis. *Biochemistry* **12**, 3055-3063 (1973).

Shields R. Common disease: are causative alleles common or rare? *PLoS.Biol.* **9**, e1001009 (2011).

Siva N. 1000 Genomes project. *Nat.Biotechnol.* **26**, 256 (2008).

Smith L. M., Sanders J. Z., Kaiser R. J., Hughes P., Dodd C., Connell C. R., Heiner C., Kent S. B., Hood L. E. Fluorescence detection in automated DNA sequence analysis. *Nature* **321**, 674-679 (1986).

Soutar A. K. Rare genetic causes of autosomal dominant or recessive hypercholesterolaemia. *IUBMB.Life* **62**, 125-131 (2010).

Stark K., Esslinger U. B., Reinhard W., Petrov G., Winkler T., Komajda M., Isnard R., Charron P., Villard E., Cambien F., Tiret L., Aumont M. C., Dubourg O., Trochu J.

N., Fauchier L., Degroote P., Richter A., Maisch B., Wichter T., Zollbrecht C., Grassl M., Schunkert H., Linsel-Nitschke P., Erdmann J., Baumert J., Illig T., Klopp N., Wichmann H. E., Meisinger C., Koenig W., Lichtner P., Meitinger T., Schillert A., Konig I. R., Hetzer R., Heid I. M., Regitz-Zagrosek V., Hengstenberg C. Genetic association study identifies HSPB7 as a risk gene for idiopathic dilated cardiomyopathy. *PLoS.Genet.* **6**, e1001167 (2010).

Theis J. L., Sharpe K. M., Matsumoto M. E., Chai H. S., Nair A. A., Theis J. D., de A. M., Wieben E. D., Michels V. V., Olson T. M. Homozygosity mapping and exome sequencing reveal GATAD1 mutation in autosomal recessive dilated cardiomyopathy. *Circ.Cardiovasc.Genet.* **4**, 585-594 (2011).

Towbin J. A., Bowles N. E. The failing heart. *Nature* **415**, 227-233 (2002).

Towbin J. A., Hejtmancik J. F., Brink P., Gelb B., Zhu X. M., Chamberlain J. S., McCabe E. R., Swift M. X-linked dilated cardiomyopathy. Molecular genetic evidence of linkage to the Duchenne muscular dystrophy (dystrophin) gene at the Xp21 locus. *Circulation* **87**, 1854-1865 (1993).

Tregouet D. A., Konig I. R., Erdmann J., Munteanu A., Braund P. S., Hall A. S., Grosshennig A., Linsel-Nitschke P., Perret C., Desuremain M., Meitinger T., Wright B. J., Preuss M., Balmforth A. J., Ball S. G., Meisinger C., Germain C., Evans A., Arveiler D., Luc G., Ruidavets J. B., Morrison C., van der Harst P., Schreiber S., Neureuther K., Schafer A., Bugert P., El Mokhtari N. E., Schrezenmeir J., Stark K., Rubin D., Wichmann H. E., Hengstenberg C., Ouwehand W., Ziegler A., Tiret L., Thompson J. R., Cambien F., Schunkert H., Samani N. J. Genome-wide haplotype association study identifies the SLC22A3-LPAL2-LPA gene cluster as a risk locus for coronary artery disease. *Nat.Genet.* **41**, 283-285 (2009).

Venter J. C., Adams M. D., Myers E. W., Li P. W., Mural R. J., Sutton G. G., Smith H. O., Yandell M., Evans C. A., Holt R. A., Gocayne J. D., Amanatides P., Ballew R. M., Huson D. H., Wortman J. R., Zhang Q., Kodira C. D., Zheng X. H., Chen L., Skupski M., Subramanian G., Thomas P. D., Zhang J., Gabor Miklos G. L., Nelson C., Broder S., Clark A. G., Nadeau J., McKusick V. A., Zinder N., Levine A. J., Roberts R. J., Simon M., Slayman C., Hunkapiller M., Bolanos R., Delcher A., Dew I., Fasulo D., Flanigan M., Florea L., Halpern A., Hannenhalli S., Kravitz S., Levy S., Mobarry C., Reinert K., Remington K., Abu-Threideh J., Beasley E., Biddick K., Bonazzi V., Brandon R., Cargill M., Chandramouliswaran I., Charlab R., Chaturvedi K., Deng Z., Di F., V, Dunn P., Eilbeck K., Evangelista C., Gabrielian A. E., Gan W., Ge W., Gong F., Gu Z., Guan P., Heiman T. J., Higgins M. E., Ji R. R., Ke Z., Ketchum K. A., Lai Z., Lei Y., Li Z., Li J., Liang Y., Lin X., Lu F., Merkulov G. V., Milshina N., Moore H. M., Naik A. K., Narayan V. A., Neelam B., Nuskern D., Rusch D. B., Salzberg S., Shao W., Shue B., Sun J., Wang Z., Wang A., Wang X., Wang J., Wei M., Wides R., Xiao C., Yan C., Yao A., Ye J., Zhan M., Zhang W., Zhang H., Zhao Q., Zheng L., Zhong F., Zhong W., Zhu S., Zhao S., Gilbert D., Baumhueter S., Spier G., Carter C., Cravchik A., Woodage T., Ali F., An H., Awe A., Baldwin D., Baden H., Barnstead M., Barrow I., Beeson K., Busam D., Carver A., Center A., Cheng M. L., Curry L., Danaher S., Davenport L., Desilets R., Dietz S., Dodson K., Doup L., Ferriera S., Garg N., Gluecksmann A., Hart B., Haynes J., Haynes C., Heiner C., Hladun S., Hostin D., Houck J., Howland T., Ibegwam C.,



- Johnson J., Kalush F., Kline L., Koduru S., Love A., Mann F., May D., McCawley S., McIntosh T., McMullen I., Moy M., Moy L., Murphy B., Nelson K., Pfannkoch C., Pratts E., Puri V., Qureshi H., Reardon M., Rodriguez R., Rogers Y. H., Romblad D., Ruhfel B., Scott R., Sitter C., Smallwood M., Stewart E., Strong R., Suh E., Thomas R., Tint N. N., Tse S., Vech C., Wang G., Wetter J., Williams S., Williams M., Windsor S., Winn-Deen E., Wolfe K., Zaveri J., Zaveri K., Abril J. F., Guigo R., Campbell M. J., Sjolander K. V., Karlak B., Kejariwal A., Mi H., Lazareva B., Hatton T., Narechania A., Diemer K., Muruganujan A., Guo N., Sato S., Bafna V., Istrail S., Lippert R., Schwartz R., Walenz B., Yooseph S., Allen D., Basu A., Baxendale J., Blick L., Caminha M., Carnes-Stine J., Caulk P., Chiang Y. H., Coyne M., Dahlke C., Mays A., Dombroski M., Donnelly M., Ely D., Esparham S., Fosler C., Gire H., Glanowski S., Glasser K., Glodek A., Gorokhov M., Graham K., Gropman B., Harris M., Heil J., Henderson S., Hoover J., Jennings D., Jordan C., Jordan J., Kasha J., Kagan L., Kraft C., Levitsky A., Lewis M., Liu X., Lopez J., Ma D., Majoros W., McDaniel J., Murphy S., Newman M., Nguyen T., Nguyen N., Nodell M. The sequence of the human genome. *Science* **291**, 1304-1351 (2001).
- Villard E., Perret C., Gary F., Proust C., Dilanian G., Hengstenberg C., Ruppert V., Arbustini E., Wichter T., Germain M., Dubourg O., Tavazzi L., Aumont M. C., Degroote P., Fauchier L., Trochu J. N., Gibelin P., Aupetit J. F., Stark K., Erdmann J., Hetzer R., Roberts A. M., Barton P. J., Regitz-Zagrosek V., Aslam U., Duboscq-Bidot L., Meyborg M., Maisch B., Madeira H., Waldenstrom A., Galve E., Cleland J. G., Dorent R., Roizes G., Zeller T., Blankenberg S., Goodall A. H., Cook S., Tregouet D. A., Tiret L., Isnard R., Komajda M., Charron P., Cambien F. A genome-wide association study identifies two loci associated with heart failure due to dilated cardiomyopathy. *Eur.Heart J.* **32**, 1065-1076 (2011).
- Vinall L. E., Hill A. S., Pigny P., Pratt W. S., Toribara N., Gum J. R., Kim Y. S., Porchet N., Aubert J. P., Swallow D. M. Variable number tandem repeat polymorphism of the mucin genes located in the complex on 11p15.5. *Hum.Genet.* **102**, 357-366 (1998).
- Vossen R. H., Aten E., Roos A., den Dunnen J. T. High-resolution melting analysis (HRMA): more than just sequence variant screening. *Hum.Mutat.* **30**, 860-866 (2009).
- Waldmüller S., Erdmann J., Binner P., Gelbrich G., Pankuweit S., Geier C., Timmermann B., Haremza J., Perrot A., Scheer S., Wachter R., Schulze-Waltrup N., Dermintzoglou A., Schonberger J., Zeh W., Jurmann B., Brodherr T., Borgel J., Farr M., Milting H., Blankenfeldt W., Reinhardt R., Ozcelik C., Osterziel K. J., Loeffler M., Maisch B., Regitz-Zagrosek V., Schunkert H., Scheffold T. Novel correlations between the genotype and the phenotype of hypertrophic and dilated cardiomyopathy: results from the German Competence Network Heart Failure. *Eur.J.Heart Fail.* **13**, 1185-1192 (2011).
- Wang L., Fan C., Topol S. E., Topol E. J., Wang Q. Mutation of MEF2A in an inherited disorder with features of coronary artery disease. *Science* **302**, 1578-1581 (2003).

- Weimer J. M., Chattopadhyay S., Custer A. W., Pearce D. A. Elevation of Hook1 in a disease model of Batten disease does not affect a novel interaction between Ankyrin G and Hook1. *Biochem.Biophys.Res.Commun.* **330**, 1176-1181 (2005).
- Weng L., Kavaslar N., Ustaszewska A., Doelle H., Schackwitz W., Hebert S., Cohen J. C., McPherson R., Pennacchio L. A. Lack of MEF2A mutations in coronary artery disease. *J.Clin.Invest* **115**, 1016-1020 (2005).
- Wible B. A., De B. M., Majumder K., Taglialatela M., Brown A. M. Cloning and functional expression of an inwardly rectifying K<sup>+</sup> channel from human atrium. *Circ.Res.* **76**, 343-350 (1995).
- Wiesner G., Grimm J., Bittner E. [Incidence of myocardial infarct in Germany: prevalence, incidence trends, East-West comparison]. *Gesundheitswesen* **61 Spec No**, S72-S78 (1999).
- Willis M. N., Liu Y., Biterova E. I., Simpson M. A., Kim H., Lee J., Barycki J. J. Enzymatic defects underlying hereditary glutamate cysteine ligase deficiency are mitigated by association of the catalytic and regulatory subunits. *Biochemistry* **50**, 6508-6517 (2011).
- Wilson P. W., Myers R. H., Larson M. G., Ordovas J. M., Wolf P. A., Schaefer E. J. Apolipoprotein E alleles, dyslipidemia, and coronary heart disease. The Framingham Offspring Study. *JAMA* **272**, 1666-1671 (1994).
- WTCCC Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* **447**, 661-678 (2007).
- Zabel U., Weeger M., La M., Schmidt H. H. Human soluble guanylate cyclase: functional expression and revised isoenzyme family. *Biochem.J.* **335 ( Pt 1)**, 51-57 (1998).

## 8. Anhang

### 8.1. Klinische Charakteristika der Familien dieser Arbeit

Tabelle 40: Charakterisierung der Probanden der DCM-Familie 16725 (vgl. Abbildung 23)

ID	DCM (1=ja, 2=nein)	Alter, (Jahre) <sup>1</sup>	Bluthochdruck <sup>2</sup> (1=ja, 2=nein)	Herz- frequenz, (bpm)	QTc-Zeit, (ms)	EF nach Simpson, (%)	NYHA- Klasse
29091	2	79	2	53	408	69	1
27972	1	69	2	75	393	27	2
29134	1	74	1	60	418	69	1
28001	1	49	2	106	406	49	
28051	1	68	2	76	441	65	1
28050	2	66	2	84	446	62	1
28360	1	57	2	96	432		2
28118	1	48	2	71	394	37	3
29313	2	53	1	91		66	1
29314	2	57	2	88		52	1
28119	1	69	2	104	419	27	2
28935	2	71	2	100	427	64	3
28936	2	65	1	75	394	66	2
28361	2	36	2	72	DM	58	1
27973	1	38	2	61	415	49	3
28003	2	19	2	80	403	59	1
28002	1	25	2	68	400	41 (49)	1
28117	2	34	2	76	397	60.9	1
28934	2	31	1	69	373	64	1
27974	1	21	1	67	398	51 (63)	1
29135	2	17	2	85	400	69	1

bpm, *beats per minute*; QTc, QT Zeit bei 60 bpm; EF, Ejektionsfraktion aus Echokardiographie nach Simpson; TM, *time motion*-Methode; <sup>1</sup> Bei Betroffenen Diagnosealter, bei Unbetroffenen Einschussalter; <sup>2</sup> Definiert als Blutdruck  $\geq 140/90$  mm Hg oder antihypertensive Behandlung

**Tabelle 41: Charakterisierung der Probanden der DCM-Familie 009-1229 (vgl. Abbildung 27)**

ID	DCM (1=ja, 2=nein)	Alter, (Jahre) <sup>1</sup>	Bluthochdruck <sup>2</sup> (1=ja, 2=nein)	Herz- frequenz, (bpm)	QTc-Zeit, (ms)	EF nach Simpson, (%)	NYHA- Klasse
01	1	68	1	80	p.t.	p.t.	2
02	1	41	1	56	468	66	1
03	1	16	2	74	446	45	1
04	2	12	2		422		1
05	2	44	2	66	432	58	1

bpm, *beats per minute*; QTc, QT Zeit bei 60 bpm; EF, Ejektionsfraktion aus Echokardiographie nach Simpson; p.t., *post transplantation*; <sup>1</sup> Alter beim Studien-Einschluss; <sup>2</sup> Definiert als Blutdruck  $\geq 140/90$  mm Hg oder antihypertensive Behandlung

**Tabelle 42: Charakterisierung der Probanden der DCM-Familie 9800 (vgl. Abbildung 21)**

ID	DCM (1=ja, 2=nein)	Alter, (Jahre) <sup>1</sup>	Bluthochdruck <sup>2</sup> (1=ja, 2=nein)	Herzfrequenz, (bpm)	EF nach Simpson, (%)	NYHA- Klasse
20077	2					
20060	2	84	2	80	65	1
18150	1	66	2	77	43,5	1
19436	2					
19184	1	64	2	71	42	2
18072	1	58	2	68	40	2-3
19236	2					
17200	1	55	2	65	35,9	2
19183	2					
19239	1	33	2	67	31,7	1
19238	2					
17306	1	55	2	65	50	1
19754	2					
19208	1	58	2	60	30	1
23882	2	46	2	55	58 TM	1
19160	1	33	2	63	41	1
19437	2	42	2	60	79,2	1
19627	2	39	2	63	61	1
23691	2	42	2	80	59	1
22557	2	28	2	64	64	1
18288	1	39	2	90	37,3	2
18703	1	35	2	115	35	2-3
18254	2	35	2	72	57	1
18073	2	28	2	60	56,3	1
18423	2	30	2	67	54	1
16719	2	34	2	66	61	1
16710	2	36	2	66	68,5	1
18648	2	33	2	86	57,8	1
19237	1	17	2	60	39,2	1
20177	2	29	2	67	68	1
17608	2	20	2	57	67	1
17671	1	7	2	64	22	1
23489	1	37	2	69	58 TM	1
17586	2	36	2	65	66,4	1
19334	2	20	2	65	64	1
25600	2	20	2	53	71	1
24903	1	19	2	78	46	1

bpm, *beats per minute*; EF, Ejektionsfraktion aus Echokardiographie nach Simpson; TM, *time motion*-Methode; <sup>1</sup> Bei Betroffenen Diagnosealter, bei Unbetroffenen Einschlussalter; <sup>2</sup> Definiert als Blutdruck  $\geq 140/90$  mm Hg oder antihypertensive Behandlung

**Tabelle 43: Charakterisierung der Probanden der HCM-Familie 757 (vgl. Abbildung 32)**

ID	HCM (1=ja, 2=nein)	Alter, (Jahre) <sup>1</sup>	Bluthochdruck <sup>2</sup> (1=ja, 2=nein)	Herzfrequenz, (bpm)	EF nach Simpson, (%)	NYHA- Klasse
2417	1	31	1	68		3
2496	1	38	2	48	67	3
1768	1	37	2	62	72	1
2453	2	30	2	75		1
1467	1	71	2	71		2
1834	2	32	2	68	70	1
2418	1	28	2	60		1
2447	2	34	2	62		1
2497	1	6	2	62	76	2
2661	2	28	2	60		1
1767	1	12	2	105	77	1
1769	1	15	2	64	72	1
1833	2	11	2	80		1
2419	1	18	2	50		2
2420	2	21	2	60		2
2454	2	12	2	68		2

bpm, *beats per minute*; EF, Ejektionsfraktion aus Echokardiographie nach Simpson; <sup>1</sup> Bei Betroffenen Diagnosealter, bei Unbetroffenen Einschlussalter; <sup>2</sup> Definiert als Blutdruck  $\geq 140/90$  mm Hg oder antihypertensive Behandlung

**Tabelle 44: Charakterisierung der Probanden der HI-Großfamilie 2001 mit DNA für genetische Analysen (vgl. Abbildung 17/40)**

ID	Einschluss Alter	BMI (kg/m <sup>2</sup> ) zum Einschluss	Follow-up Alter	HI- Erstereignis- Alter	KHK- Erstereignis- Alter	HI- oder KHK- Ersteralter	Bluthoch- druck <sup>1</sup> (1=ja, 2=nein)	Diabetes <sup>2</sup> (1=ja, 2=nein)	Hypercholesterinämie <sup>3</sup> (1=ja, 2=nein)	HDL- Cholesterin (mg/dL) zum Studien- einschluss	Triglyceride (mg/dL)um Studien- einschluss	Rauchen <sup>4</sup> (1=ja, 2=nein)
2001588	65,6	27,9	76,1	58		58	1	1	2	45	129	2
2001601	59,7	27,8	59,7		55	55	1	2	1	40	169	1
2001589	58,7	21,7	62,0				1	2		34	389	2
2001590	63,1	30,5	72,1				1	2	1	33	228	2
2001596	61,1	30,0	64,4	55		55		1		32	322	2
2001594	50,4	26,7	62,0	33		33	1	2	2	37	132	1
2001593	73,0	28,0	81,9		82	82 *	1	2	2	44	106	2
2001513	61,5	28,1	67,0		59	59	1	2	1	40	495	1
2001516	56,8	23,3	66,8	50		50	1	2	1	37	134	1
2001517	59,3	25,9	68,3		66	66	1	2	2	57	352	2
2001552	65,6	25,2	71,1	53		53	1	2	1	53	192	2
2001511	53,3	26,4	64,0	52		52	1	2	1	57	297	2
2001529	58,9	27,8	65,0	57		57	1	1	1	32	381	1
2001577	63,0	27,7	73,2				1	2	1	44	162	1
2001525	60,8	25,2	69,7				1	2	1	49	109	2
2001521	53,6	29,4	64,9		61	61	1	2	1	49	261	1
2001567	59,6	29,1	65,1		59	59	1	1	1	42	478	1
2001628	61,8	26,0	69,0				1	2	2	56	77	1
2001580	63,3	28,4	72,2				1	2	2	45	137	1
2001544	62,3	27,0	72,3	47		47	1	2	2	48	217	2
2001547	56,4	31,6	65,4		58	58	1	2		27	217	2
2001549	68,9	22,5	68,9	67		67	1	2	1	44	198	2
2001603	65,2		75,2					2	1	39	135	

<sup>1</sup> Definiert als Blutdruck  $\geq 140/90$  mm Hg oder antihypertensive Behandlung; <sup>2</sup> Definiert als anamnestischer Diabetes mellitus, HbA1c  $> 6,5$  % oder Einnahme von antidiabetischen Medikamenten; <sup>3</sup> Definiert als LDL  $\geq 160$  mg/dL oder lipidsenkende Therapie; <sup>4</sup> Früheres oder aktuelles Rauchen zum Studien-Einschluss; \* ID2001593 als unbetroffen, weil KHK-Erstereignis im hohen Alter von 82 Jahren

## 8.2. Klinische Charakteristika der Screening-Kollektive

**Tabelle 45: Charakterisierung des Kollektivs für das Screening von *GUCY1A3***

Variable	HI-Patienten nicht familiär	Gesunde Individuen	HI-Patienten aus Großfamilien
	(n = 483)	(n = 562)	(n = 81)
Alter, Jahre (Altersbereich) <sup>a</sup>	54,7 ± 8,4 (29 - 75)	63,6 ± 4,6 (56 - 80)	62,5 ± 7,9 (43 - 79)
Geschlecht, % Männer (n)	84,7 (409)	29,4 (165)	69,1 (56)
HI % (n)	100 (483)	-	93,8 (76)
Alter bei erstem HI, Jahre	43,2 (13-52)	-	54,0 (33-73)
Hypercholesterinämie <sup>b</sup> , % (n)	79,5 (384)	48,5 (271)	80,0 (64)
Lipid senkende Medikation, %	64,9 (303)	17,1 (74)	60,5 (46)
LDL-C, mg/dl	146,9 ± 41,7	152,8 ± 33,5	157,1 ± 41,4
HDL-C, mg/dl	48,7 ± 12,9	61,4 ± 14,8	51,8 ± 13,5
Hypertension <sup>c</sup> , % (n)	92,0 (439)	70,4 (375)	93,8 (75)
Antihypertensive Therapie, %	87,6 (409)	57,5 (249)	90,8 (69)
Systolischer Blutdruck, mmHg	134,6 ± 17,1	136,8 ± 17,7	147,7 ± 21,8
Diastolischer Blutdruck, mmHg	82,2 ± 9,7	82,7 ± 9,1	84,7 ± 10,3
Typ 2 Diabetes <sup>d</sup> , % (n)	13,5 (65)	5,9 (33)	11,1 (9)
Raucher <sup>e</sup> , % (n)	79,0 (380)	37,1 (208)	71,3 (57)
BMI, kg/m <sup>2</sup>	27,6 ± 3,8	26,7 ± 3,7	27,1 ± 3,2

Werte sind angegeben ± der Standardabweichung; n, Anzahl; LDL-C, *low-density* Lipoprotein Cholesterin; HDL-C, *high-density* Lipoprotein Cholesterin; BMI, *body mass index*

<sup>a</sup> Bei Einschluss in die Studie

<sup>b</sup> Definiert als LDL-C ≥ 160 mg/dL oder Einnahme von Lipid senkender Medikation

<sup>c</sup> Definiert als Blutdruck ≥ 140/90 mmHg oder anihypertensive Therapie

<sup>d</sup> Definiert als Krankheitsgeschichte mit Diabetes mellitus oder Einnahme von antidiabetischer Medikation

<sup>e</sup> Früheres oder aktuelles Rauchen zum Studien-Einschluss

**Tabelle 46: Charakterisierung des Kollektivs für die Sequenzierung dreier Gene**

Variable	HI-Patienten aus Großfamilien
	(n = 48)
Alter, Jahre (Altersbereich) <sup>a</sup>	60,5 ± 8,2 (38 - 72)
Geschlecht, % Männer (n)	63,8 (30)
HI % (n)	100 (48)
Alter bei erstem HI, Jahre	50,8 ± 8,3 (35-67)

<sup>a</sup> Bei Einschluss in die Studie



### 8.3. Oligonukleotide

**Tabelle 47: Oligonukleotide für Ko-Segregation in Familie 2001**

Produkt	<i>forward</i> Primer	<i>reverse</i> Primer	Größe in bp	T <sub>A</sub>	Zyklen
<i>GUCY1A3</i> Exon 7	tgtgattcaggagtccagt	cgacactccactccggtt	338	57°C	40
<i>ETFDH</i> Exon 2	ttgtgcagcatatcagtg	agctacgatggctggaagc	409	57°C	40
<i>GCLC</i> Exon 10	gcagtgctcgaggataaa	ctgcactgggtaaataagtc	454	57°C	40
<i>CCT7</i> Exon 12	cgggttctagcagtgctt	caagccaaggaggtaagtg	376	59°C	40

**Tabelle 48: Oligonukleotide für IDAHO-Screening *GUCY1A3***

Produkt	<i>forward</i> Primer	<i>reverse</i> Primer	Größe in bp	T <sub>A</sub>	Zyklen
<i>GUCY1A3</i> Exon 4	tttcatccgtgcataaagtca	gccttgaaagtgagaaggaagt	525	63°C	40
<i>GUCY1A3</i> Exon 5	aaatcataacattatcattcaagc	aactgctggccattagtagca	277	61°C	40
<i>GUCY1A3</i> Exon 6	cgctgtgcctggcctaat	catactgggagtcaccaacaaa	274	63°C	40
<i>GUCY1A3</i> Exon 7_1	cagctttaccaaggtggcat	gacagggatggcttgggtg	480	64°C	40
<i>GUCY1A3</i> Exon 7_2	aaacggaagtgaagtgtcg	ttgttctgtcttaaatatgaaat	467	58°C	40
<i>GUCY1A3</i> Exon 8_1	ttgtctcagaaagatgtgggaa	gagcaacctcacagggaat	365	61°C	40
<i>GUCY1A3</i> Exon 8_2	gatggcctgaagaagaggct	taggcacagtgttcttccc	372	61°C	40
<i>GUCY1A3</i> Exon 9	gcagtgctcaatgtaatgg	gtcctctcaatgattctgtaagc	552	63°C	40
<i>GUCY1A3</i> Exon 10	tcaaacgacactgacgagtagg	gggcaaatggcaattgta	288	60°C	40
<i>GUCY1A3</i> Exon 11	tgggaagatgtctgtaatctcg	ggagaaagattgtgactggc	494	62°C	40

**Tabelle 49: Oligonukleotide für Sequenzierung von *ETFDH*, *CCT7* und *GCLC***

Produkt	<i>forward</i> Primer	<i>reverse</i> Primer	Größe in bp	T <sub>A</sub>	Zyklen
<i>ETFDH</i> Exon 1	aacgccgtgaagcaagag	tgaccttgcctagatagcc	331	63°C	40
<i>ETFDH</i> Exon 2	gcctgacatgagctaaattgaa	tgacgatgtgacttgtaa	500	60°C	40
<i>ETFDH</i> Exon 3	aacacagggagaattccataaa	ttgggaacaactactgaaatacaa	416	60°C	40
<i>ETFDH</i> Exon 4	gagaaaggattcacagtgacaagtt	ctagtctgcttcccacca	407	60°C	40
<i>ETFDH</i> Exon 5	aaagtgtgaccatcaatgtagca	gggtacaacatgaggaattcaa	428	60°C	40
<i>ETFDH</i> Exon 6	gaaacctaaagctgtttactgttt	ggttcactttgatgccac	237	60°C	40
<i>ETFDH</i> Exon 7	tttgactgtctcttctacatctg	tttgatcagtaatatcttctagccc	440	63°C	40
<i>ETFDH</i> Exon 8	aaataagacattaaacctggcaa	gacacagcagcataagcac	395	60°C	40
<i>ETFDH</i> Exon 9	ttactgcacatagtgctccaaa	ccactgtgccagcctaata	479	60°C	40
<i>ETFDH</i> Exon 10	ctgatttgagatattctgtgttct	gtcaattgaaatacacataaccag	304	60°C	40
<i>ETFDH</i> Exon 11	gagcaagacctgtctcgaa	tcactttctctaattaacagatcca	382	62°C	40
<i>ETFDH</i> Exon 12	gggctagtcatattcttgggtg	caatttctaaggaatggaagga	444	57°C	40
<i>ETFDH</i> Exon 13	ttctgtggctactcttccctaat	tggtactgaatctctgaagaactg	498	60°C	40
<i>CCT7</i> Exon 1	taagaagcacgaccgaaatg	ggctctacagagcaagccc	350	60°C	40
<i>CCT7</i> Exon 2	ttaagggagagatgggagcat	ccaacctatgctagtcccaa	324	65°C	40
<i>CCT7</i> Exon 3	aggctccactgaagacagc	aggagcaagcaagcatgtgt	225	60°C	40
<i>CCT7</i> Exon 4	cagggctttattccctgaac	cccttctaagagaaatgtgaa	245	60°C	40
<i>CCT7</i> Exon 5	gaggtccacttcagcagtttg	taacgaaggtcccgtctct	170	63°C	40
<i>CCT7</i> Exon 6	gaatcagataagagggatttggtt	ccctctgccatgaagctg	308	60°C	40
<i>CCT7</i> Exon 7	tggataagcactagccagca	ggctcagaagctgccc	304	63°C	40
<i>CCT7</i> Exon 8	ccctaccgtatgtcaaccttc	aaaccttcagggccac	302	60°C	40
<i>CCT7</i> Exon 9	attggctttgagacgaggtg	agatggcagaagaccagacc	238	60°C	40
<i>CCT7</i> Exon 10	cagtaggatgggtggcagaca	caccacacaaatccaaggtct	270	60°C	40
<i>CCT7</i> Exon 11	agcatggaagcttgaccag	gcttgagtaagcacagctcaca	344	65°C	40
<i>CCT7</i> Exon 12	tggagtaccagtttacaagagga	ttggacaagtgcattgaaagc	522	60°C	40
<i>GCLC</i> Exon 1	cggctgaggtccgtctc	gcgtagggcaagacaaagg	304	60°C	40
<i>GCLC</i> Exon 2	cagacttcagttccctgc	ggctatcctgcctccaataa	297	60°C	40
<i>GCLC</i> Exon 3	gaaggactcttagcttcattatacc	gattgtaagtggaatgcc	401	60°C	40
<i>GCLC</i> Exon 4	aatgaggtgtctttggcagg	tgaacaggtgcacatctgctat	366	60°C	40
<i>GCLC</i> Exon 5-6	ttgaggtatgattcattccacttac	ccagctaccattcaacctca	517	60°C	40
<i>GCLC</i> Exon 7	ctctgacttactgctggg	tctgcacatgctatcttattcata	208	60°C	40
<i>GCLC</i> Exon 8	ggtacatgtgctctgtatggg	gcagggagacatcctgaaat	333	60°C	40
<i>GCLC</i> Exon 9	Keine Amplifikation				
<i>GCLC</i> Exon 10	aagaggaagtggtatctctcgc	aaacagaggattccagactctca	236	60°C	40
<i>GCLC</i> Exon 112	tttggtatggacagtgagcag	agggctgatggagaacatt	748	60°C	40
<i>GCLC</i> Exon 13-15	tgcaaaggtgaggaaggtg	tgtgaaactgtcatgaatacag	665	62°C	40
<i>GCLC</i> Exon 16	tccagaattgtttatgctga	ttcaataaatcaggtcccagg	710	62°C	40

**Tabelle 50: Oligonukleotide für Ko-Segregation in Familie 009-1229 und 757**

Produkt	<i>forward</i> Primer	<i>reverse</i> Primer	Größe in bp	T <sub>A</sub>	Zyklen
<i>GPR98</i> Exon 17	ctcgtgtagtgagggttcagg	gggtcaacacatacattcttcacc	440	60°C	40
<i>ZG16</i> Exon 4	gtaggatcttcagggttcac	gggtaaacaatcccagtgacag	403	60°C	40
<i>INTS1</i> Exon 19	cggcttctcgtctcagctct	ctgctccttcagcctctgtc	466	60°C	40
<i>CYFIP1</i> Exon 1	cctcctgtctgtccgtgtct	tgagggactctagcatggttc	455	60°C	40
<i>PINX1</i> Exon 7	gagggaaagcccagagagg	gggtgaactctgctgtgactt	370	60°C	40
<i>WHAMM</i> Exon 9	tgggttcggtagaaaggatg	ttccacactagaggacgtg	591	60°C	11/33
<i>ZNF98</i> Exon 4	gtgtggctcttgcattgg	tggtgagctaagtgtgaaagca	581	60°C	11/33
<i>NPPA</i> Exon 2	acgcagacctgatggatttc	ctccaatcctgtccatcctg	440	60°C	40
<i>GJB4</i> Exon 2	tcaatcgaccagcattaag	tgctcaggtgtcgtacagg	491	60°C	40
<i>ZNF598</i> Exon 9	gcctccaggaggaggact	attaacgacgacggcaactg	465	60°C	40
<i>HOOK1</i> Exon 7	cctcactaatatgcttgaattgtgt	gaccaacaccactgccattc	423	60°C	40

## 8.4. Parameter von SAMtools

**Tabelle 51: Parameter von SAMtools für Filterung der Varianten**

<b>SAMtools v.0.1.7 varFilter</b>	<b>SAMtools v.0.1.16</b>
-Q minimum RMS mapping quality for SNPs [25]	-Q minimum value of the QUAL field [10]
-q minimum RMS mapping quality for gaps [10]	-q minimum RMS mapping quality for SNPs [10]
-d minimum read depth [3]	-d minimum read depth [2]
-D maximum read depth [9999]	-D maximum read depth [1000000]
-G min indel score for nearby SNP filtering [25]	-w SNP within INT bp around a gap to be filtered [10]
-w SNP within INT bp around a gap to be filtered [10]	-W window size for filtering adjacent gaps [3]
-W window size for filtering dense SNPs [10]	
-N max number of SNPs in a window [2]	
-l window size for filtering adjacent gaps [30]	

## 8.5. Stammbaum HI-Großfamilie 2001

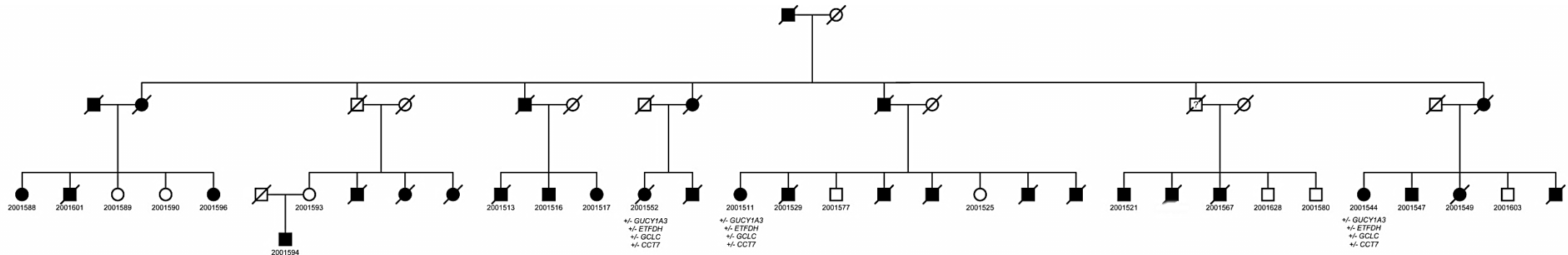


Abbildung 40: Stammbaum HI-Großfamilie 2001. Quadrat: Mann, Kreis: Frau, ausgefüllte Symbole: betroffen, offene Symbole: nicht betroffen, durchgestrichene Symbole: verstorben, ?: Phänotyp unbekannt.

## 8.6. Anzahl detektierter Varianten

Tabelle 52: Anzahl detektierter Varianten in DCM-Familie 9800

ID	Typ	captured/ ohne rs- Nummer	ohne rs- Nummer in %	kodierend/ ohne rs- Nummer	ohne rs- Nummer in %	synonym/ ohne rs- Nummer	ohne rs- Nummer in %	nicht-synonym, Indel, Spleiß- Variante/ ohne rs-Nummer	ohne rs- Nummer in %
17200	Alle	64826/ 3794	5,9	22124/ 1072	4,9	10070/ 314	3,1	12053/ 778	6,5
	Homozygot	22707/ 299	1,3	7977/ 51	0,6	3709/ 7	0,2	4267/ 44	1,0
	Heterozygot	42120/ 3495	8,3	14147/ 1021	7,2	6361/ 307	4,8	7786/ 734	9,4

ID	Typ	captured/ ohne rs- Nummer	ohne rs- Nummer in %	kodierend/ ohne rs- Nummer	ohne rs- Nummer in %	synonym/ ohne rs- Nummer	ohne rs- Nummer in %	nicht-synonym, Indel, Spleiß- Variante/ ohne rs-Nummer	ohne rs- Nummer in %
17200 modif.	Alle	71287/ 4139	5,8	23034/ 1177	5,1	10413/ 355	3,4	12621/ 724	5,7
	Homozygot	24479/ 323	1,3	8198/ 62	0,8	3793/ 12	0,3	4405/ 109	2,5
	Heterozygot	46808/ 3816	8,2	14836/ 1115	7,5	6620/ 343	5,2	8216/ 692	8,4

ID	Typ	captured/ ohne rs- Nummer	ohne rs- Nummer in %	kodierend/ ohne rs- Nummer	ohne rs- Nummer in %	synonym/ ohne rs- Nummer	ohne rs- Nummer in %	nicht-synonym, Indel, Spleiß- Variante/ ohne rs-Nummer	ohne rs- Nummer in %
17671	Alle	66226/ 3845	5,8	22384/ 1088	4,9	10187/ 333	3,3	12197/ 755	6,2
	Homozygot	22697/ 317	1,4	7911/ 63	0,8	3642/ 17	0,5	4269/ 46	1,1
	Heterozygot	43529/ 3528	8,1	14473/ 1025	7,1	6545/ 316	4,8	7928/ 709	8,9

ID	Typ	captured/ ohne rs- Nummer	ohne rs- Nummer in %	kodierend/ ohne rs- Nummer	ohne rs- Nummer in %	synonym/ ohne rs- Nummer	ohne rs- Nummer in %	nicht-synonym, Indel, Spleiß- Variante/ ohne rs-Nummer	ohne rs- Nummer in %
17671 modif.	Alle	72275/ 4140	5,7	23259/ 1185	5,1	10446/ 343	3,3	12813/ 842	6,6
	Homozygot	24670/ 333	1,4	8133/ 55	0,7	3704/ 15	0,4	4429/ 40	0,9
	Heterozygot	47605/ 3807	8,0	15126/ 1130	7,5	6742/ 328	4,9	8384/ 802	9,6

ID	Typ	captured/ ohne rs- Nummer	ohne rs- Nummer in %	kodierend/ ohne rs- Nummer	ohne rs- Nummer in %	synonym/ ohne rs- Nummer	ohne rs- Nummer in %	nicht-synonym, Indel, Spleiß- Variante/ ohne rs-Nummer	ohne rs- Nummer in %
18288	Alle	67614/ 4103	6,1	22734/ 1134	5,0	10388/ 307	3,0	12346/ 827	6,7
	Homozygot	23442/ 414	1,8	8231/ 79	1,0	3836/ 15	0,4	4395/ 64	1,5
	Heterozygot	44172/ 3689	8,4	14503/ 1055	7,3	6552/ 292	4,5	7951/ 763	9,6

ID	Typ	captured/ ohne rs- Nummer	ohne rs- Nummer in %	kodierend/ ohne rs- Nummer	ohne rs- Nummer in %	synonym/ ohne rs- Nummer	ohne rs- Nummer in %	nicht-synonym, Indel, Spleiß- Variante/ ohne rs-Nummer	ohne rs- Nummer in %
<b>18288 modif.</b>	Alle	72782/ 4402	6,1	23837/ 1241	5,2	10893/ 343	3,2	12944/ 898	6,9
	Homozygot	24853/ 357	1,4	8593/ 66	0,8	4010/ 15	0,4	4583/ 51	1,1
	Heterozygot	47929/ 4045	8,4	15244/ 1175	7,7	6883/ 328	4,8	8361/ 847	10,1

ID	Typ	captured/ ohne rs- Nummer	ohne rs- Nummer in %	kodierend/ ohne rs- Nummer	ohne rs- Nummer in %	synonym/ ohne rs- Nummer	ohne rs- Nummer in %	nicht-synonym, Indel, Spleiß- Variante/ ohne rs-Nummer	ohne rs- Nummer in %
<b>19208</b>	Alle	64332/ 3742	5,8	22075/ 1056	4,8	10002/ 307	3,1	12073/ 749	6,2
	Homozygot	22446/ 380	1,7	7900/ 78	1,0	3621/ 14	0,4	4279/ 64	1,5
	Heterozygot	41886/ 3362	8,0	14175/ 978	6,9	6381/ 293	4,6	7794/ 685	8,8

ID	Typ	captured/ ohne rs- Nummer	ohne rs- Nummer in %	kodierend/ ohne rs- Nummer	ohne rs- Nummer in %	synonym/ ohne rs- Nummer	ohne rs- Nummer in %	nicht-synonym, Indel, Spleiß- Variante/ ohne rs-Nummer	ohne rs- Nummer in %
<b>19208 modif.</b>	Alle	69357/ 4129	6,0	22990/ 1155	5,0	10392/ 338	3,3	12598/ 844	6,7
	Homozygot	23965/ 408	1,7	8195/ 85	1,0	3745/ 16	0,4	4450/ 69	1,6
	Heterozygot	45392/ 3721	8,2	14795/ 1070	7,2	6647/ 322	4,8	8148/ 775	9,5

ID	Typ	captured/ ohne rs- Nummer	ohne rs- Nummer in %	kodierend/ ohne rs- Nummer	ohne rs- Nummer in %	synonym/ ohne rs- Nummer	ohne rs- Nummer in %	nicht-synonym, Indel, Spleiß- Variante/ ohne rs-Nummer	ohne rs- Nummer in %
<b>19237</b>	Alle	63614/ 3840	6,0	21969/ 1044	4,8	9949/ 304	3,1	12020/ 740	6,2
	Homozygot	22482/ 412	1,8	7978/ 82	1,0	3713/ 21	0,6	4265/ 61	1,4
	Heterozygot	41132/ 3428	8,3	13991/ 962	6,9	6236/ 283	4,5	7755/ 679	8,8

ID	Typ	captured/ ohne rs- Nummer	ohne rs- Nummer in %	kodierend/ ohne rs- Nummer	ohne rs- Nummer in %	synonym/ ohne rs- Nummer	ohne rs- Nummer in %	nicht-synonym, Indel, Spleiß- Variante/ ohne rs-Nummer	ohne rs- Nummer in %
<b>19237 modif.</b>	Alle	70426/ 4112	5,8	23071/ 1151	5,0	10457/ 333	3,2	12614/ 818	6,5
	Homozygot	24577/ 383	1,6	8292/ 81	1,0	3866/ 25	0,7	4426/ 56	1,3
	Heterozygot	45849/ 3729	8,1	14779/ 1070	7,2	6591/ 308	4,7	8188/ 762	9,3

Tabelle 53: Anzahl detektierter Varianten in DCM-Familie 16725

ID	Typ	captured/ ohne rs- Nummer	ohne rs- Nummer in %	kodierend/ ohne rs- Nummer	ohne rs- Nummer in %	synonym/ ohne rs- Nummer	ohne rs- Nummer in %	nicht-synonym, Indel, Spleiß- Variante/ ohne rs-Nummer	ohne rs- Nummer in %
<b>27974</b>	Alle	54026/ 2636	4,9	19019/ 899	4,7	8594/ 255	3,0	10425/ 644	6,2
	Homozygot	21573/ 230	1,1	7241/ 54	0,7	3354/ 8	0,2	3887/ 46	1,2
	Heterozygot	32453/ 2406	7,4	11778/ 845	7,2	5240/ 247	4,7	6538/ 598	9,1

ID	Typ	captured/ ohne rs- Nummer	ohne rs- Nummer in %	kodierend/ ohne rs- Nummer	ohne rs- Nummer in %	synonym/ ohne rs- Nummer	ohne rs- Nummer in %	nicht-synonym, Indel, Spleiß- Variante/ ohne rs-Nummer	ohne rs- Nummer in %
<b>27974 modif.</b>	Alle	17077/ 1139	6,7	7137/ 500	7,0	2985/ 114	3,8	4152/ 386	9,3
	Homozygot	7295/ 82	1,1	2965/ 28	0,9	1258/ 5	0,4	1707/ 23	1,3
	Heterozygot	9782/ 1057	10,8	4172/ 472	11,3	1727/ 109	6,3	2445/ 363	14,8

ID	Typ	captured/ ohne rs- Nummer	ohne rs- Nummer in %	kodierend/ ohne rs- Nummer	ohne rs- Nummer in %	synonym/ ohne rs- Nummer	ohne rs- Nummer in %	nicht-synonym, Indel, Spleiß- Variante/ ohne rs-Nummer	ohne rs- Nummer in %
<b>28002</b>	Alle	69602/ 3778	5,4	22565/ 1092	4,8	10179/ 346	3,4	12836/ 746	6,0
	Homozygot	23791/ 349	1,5	7896/ 70	0,9	3673/ 18	0,5	4223/ 52	1,2
	Heterozygot	45811/ 3429	7,5	14669/ 1022	7,0	6506/ 328	5,0	8163/ 694	8,5

ID	Typ	captured/ ohne rs- Nummer	ohne rs- Nummer in %	kodierend/ ohne rs- Nummer	ohne rs- Nummer in %	synonym/ ohne rs- Nummer	ohne rs- Nummer in %	nicht-synonym, Indel, Spleiß- Variante/ ohne rs-Nummer	ohne rs- Nummer in %
<b>28002 modif.</b>	Alle	63514/ 3004	4,7	20997/ 1011	4,8	9412/ 326	3,5	11585/ 685	5,9
	Homozygot	24658/ 262	1,1	7767/ 55	0,7	3629/ 16	0,4	4138/ 39	0,9
	Heterozygot	38856/ 2742	7,1	13230/ 956	7,2	5783/ 310	5,4	7447/ 646	8,7

ID	Typ	captured/ ohne rs- Nummer	ohne rs- Nummer in %	kodierend/ ohne rs- Nummer	ohne rs- Nummer in %	synonym/ ohne rs- Nummer	ohne rs- Nummer in %	nicht-synonym, Indel, Spleiß- Variante/ ohne rs-Nummer	ohne rs- Nummer in %
<b>28118</b>	Alle	28297/ 1522	5,4	10524/ 555	5,3	4564/ 149	3,3	5960/ 406	6,8
	Homozygot	11866/ 115	1,0	4252/ 27	0,6	1937/ 4	0,2	2315/ 23	1,0
	Heterozygot	16431/ 1407	8,6	6272/ 528	8,4	2627/ 145	5,5	3645/ 383	10,5

ID	Typ	captured/ ohne rs- Nummer	ohne rs- Nummer in %	kodierend/ ohne rs- Nummer	ohne rs- Nummer in %	synonym/ ohne rs- Nummer	ohne rs- Nummer in %	nicht-synonym, Indel, Spleiß- Variante/ ohne rs-Nummer	ohne rs- Nummer in %
<b>28118 modif.</b>	Alle	7611/ 621	8,2	3111/ 249	8,0	1250/ 64	5,1	1861/ 185	9,9
	Homozygot	3068/ 45	1,5	1151/ 18	1,6	487/ 6	1,2	664/ 12	1,8
	Heterozygot	4543/ 576	12,7	1960/ 231	11,8	763/ 58	7,6	1197/ 173	14,5

ID	Typ	captured/ ohne rs- Nummer	ohne rs- Nummer in %	kodierend/ ohne rs- Nummer	ohne rs- Nummer in %	synonym/ ohne rs- Nummer	ohne rs- Nummer in %	nicht-synonym, Indel, Spleiß- Variante/ ohne rs-Nummer	ohne rs- Nummer in %
<b>28119</b>	Alle	46890/ 2455	5,2	16987/ 913	5,4	7563/ 250	3,3	9424/ 663	7,0
	Homozygot	18989/ 175	0,9	6547/ 35	0,5	3003/ 8	0,3	3544/ 27	0,8
	Heterozygot	27901/ 2280	8,2	10440/ 878	8,4	4560/ 242	5,3	5880/ 636	10,8

ID	Typ	captured/ ohne rs- Nummer	ohne rs- Nummer in %	kodierend/ ohne rs- Nummer	ohne rs- Nummer in %	synonym/ ohne rs- Nummer	ohne rs- Nummer in %	nicht-synonym, Indel, Spleiß- Variante/ ohne rs-Nummer	ohne rs- Nummer in %
<b>28119 modif.</b>	Alle	14618/ 1053	7,2	6000/ 486	8,1	2494/ 131	5,3	3506/ 355	10,1
	Homozygot	6047/ 63	1,0	2340/ 22	0,9	1015/ 6	0,6	1325/ 16	1,2
	Heterozygot	8571/ 990	11,6	3660/ 464	12,7	1479/ 125	8,5	2181/ 339	15,5

ID	Typ	captured/ ohne rs- Nummer	ohne rs- Nummer in %	kodierend/ ohne rs- Nummer	ohne rs- Nummer in %	synonym/ ohne rs- Nummer	ohne rs- Nummer in %	nicht-synonym, Indel, Spleiß- Variante/ ohne rs-Nummer	ohne rs- Nummer in %
<b>29134</b>	Alle	70198/ 4117	5,9	23049/ 1202	5,2	10394/ 364	3,5	12655/ 838	6,6
	Homozygot	24217/ 357	1,5	8091/ 73	0,9	3766/ 18	0,5	4325/ 55	1,3
	Heterozygot	45981/ 3760	8,2	14958/ 1129	7,5	6628/ 346	5,2	8330/ 783	9,4

ID	Typ	captured/ ohne rs- Nummer	ohne rs- Nummer in %	kodierend/ ohne rs- Nummer	ohne rs- Nummer in %	synonym/ ohne rs- Nummer	ohne rs- Nummer in %	nicht-synonym, Indel, Spleiß- Variante/ ohne rs-Nummer	ohne rs- Nummer in %
<b>29134 modif.</b>	Alle	69455/ 3656	5,3	22577/ 1222	5,4	10168/ 372	3,7	12409/ 850	6,8
	Homozygot	25911/ 330	1,3	8080/ 54	0,7	3756/ 15	0,4	4324/ 39	0,9
	Heterozygot	43544/ 3326	7,6	14497/ 1168	8,1	6412/ 357	5,6	8085/ 811	10,0



Tabelle 54: Anzahl detektierter Varianten in DCM-Familie 009-1229

ID	Typ	captured/ ohne rs- Nummer	ohne rs- Nummer in %	kodierend/ ohne rs- Nummer	ohne rs- Nummer in %	synonym/ ohne rs- Nummer	ohne rs- Nummer in %	nicht-synonym, Indel, Spleiß- Variante/ ohne rs-Nummer	ohne rs- Nummer in %
01	Alle	67523/ 3851	5,7	22635/ 1099	4,9	10253/ 340	3,3	12382/ 758	6,1
	Homozygot	23168/ 343	1,5	8152/ 66	0,8	3809/ 16	0,4	4343/ 49	1,1
	Heterozygot	44355/ 3508	7,9	14483/ 1033	7,1	6444/ 324	5,0	8039/ 709	8,8

ID	Typ	captured/ ohne rs- Nummer	ohne rs- Nummer in %	kodierend/ ohne rs- Nummer	ohne rs- Nummer in %	synonym/ ohne rs- Nummer	ohne rs- Nummer in %	nicht-synonym, Indel, Spleiß- Variante/ ohne rs-Nummer	ohne rs- Nummer in %
01 modif.	Alle	71260/ 4289	6,0	23547/ 1229	5,2	10610/ 385	3,6	12937/ 844	6,5
	Homozygot	24292/ 386	1,6	8436/ 75	0,9	3922/ 16	0,4	4514/ 59	1,3
	Heterozygot	46968/ 3903	8,3	15111/ 1154	7,6	6688/ 369	5,5	8423/ 785	9,3

ID	Typ	captured/ ohne rs- Nummer	ohne rs- Nummer in %	kodierend/ ohne rs- Nummer	ohne rs- Nummer in %	synonym/ ohne rs- Nummer	ohne rs- Nummer in %	nicht-synonym, Indel, Spleiß- Variante/ ohne rs-Nummer	ohne rs- Nummer in %
03	Alle	67161/ 3882	5,8	22693/ 1065	4,7	10276/ 315	3,1	12417/ 750	6,0
	Homozygot	23091/ 352	1,5	8057/ 68	0,8	3776/ 14	0,4	4281/ 54	1,3
	Heterozygot	44070/ 3530	8,0	14636/ 997	6,8	6500/ 301	4,6	8136/ 696	8,6

ID	Typ	captured/ ohne rs- Nummer	ohne rs- Nummer in %	kodierend/ ohne rs- Nummer	ohne rs- Nummer in %	synonym/ ohne rs- Nummer	ohne rs- Nummer in %	nicht-synonym, Indel, Spleiß- Variante/ ohne rs-Nummer	ohne rs- Nummer in %
03 modif.	Alle	70516/ 4102	5,8	23359/ 1142	4,9	10565/ 331	3,1	12794/ 811	6,3
	Homozygot	24005/ 364	1,5	8228/ 71	0,9	3858/ 12	0,3	4370/ 59	1,4
	Heterozygot	46511/ 3738	8,0	15131/ 1071	7,1	6707/ 319	4,8	8424/ 752	8,9

Tabelle 55: Anzahl detektierter Varianten in HCM-Familie 757

ID	Typ	captured/ ohne rs- Nummer	ohne rs- Nummer in %	kodierend/ ohne rs- Nummer	ohne rs- Nummer in %	synonym/ ohne rs- Nummer	ohne rs- Nummer in %	nicht-synonym, Indel, Spleiß- Variante/ ohne rs-Nummer	ohne rs- Nummer in %
1769	Alle	80541/ 4183	5,2	23074/ 1078	4,7	10460/ 304	2,9	12614/ 774	6,1
	Homozygot	30728/ 406	1,3	8385/ 67	0,8	3922/ 10	0,3	4463/ 57	1,3
	Heterozygot	49813/ 3777	7,6	14689/ 1011	6,9	6538/ 294	4,5	8151/ 717	8,8

ID	Typ	captured/ ohne rs- Nummer	ohne rs- Nummer in %	kodierend/ ohne rs- Nummer	ohne rs- Nummer in %	synonym/ ohne rs- Nummer	ohne rs- Nummer in %	nicht-synonym, Indel, Spleiß- Variante/ ohne rs-Nummer	ohne rs- Nummer in %
2419	Alle	78989/ 3938	5,0	22464/ 1024	4,6	10080/ 294	2,9	12384/ 730	5,9
	Homozygot	30516/ 412	1,4	8348/ 73	0,9	3840/ 14	0,4	4508/ 59	1,3
	Heterozygot	48473/ 3526	7,3	14116/ 951	6,7	6240/ 280	4,5	7876/ 671	8,5

ID	Typ	captured/ ohne rs- Nummer	ohne rs- Nummer in %	kodierend/ ohne rs- Nummer	ohne rs- Nummer in %	synonym/ ohne rs- Nummer	ohne rs- Nummer in %	nicht-synonym, Indel, Spleiß- Variante/ ohne rs-Nummer	ohne rs- Nummer in %
2497	Alle	78845/ 4137	5,2	22465/ 1071	4,8	10207/ 323	3,2	12258/ 748	6,1
	Homozygot	30983/ 459	1,5	8427/ 88	1,0	3828/ 16	0,4	4599/ 72	1,6
	Heterozygot	47862/ 3678	7,7	14038/ 983	7,0	6379/ 307	4,8	7659/ 676	8,8

---

## Danksagung

An dieser Stelle möchte ich mich bei Herrn Prof. Dr. Christian Hengstenberg für die interessante und spannende Themenstellung und die Korrektur und Begutachtung dieser Arbeit herzlich bedanken. Frau Prof. Dr. Iris Heid danke ich für die Erstellung des Zweitgutachtens.

Bei PD Dr. Klaus Stark bedanke ich mich für die Betreuung, Hilfe, Unterstützung und Korrektur der Doktorarbeit.

Ich danke meinen Kollegen Maya Fürstenau-Sharp, Nico Jensch, Melanie Klingenstein, Michaela Schmidt, Josef Simon, Martina Zimmermann und Christa Zollbrecht für den angenehmen gemeinsamen Laboralltag. Besonders möchte ich mich bei Christa Zollbrecht für eine tolle Freundschaft und eine schöne Zeit bedanken. Außerdem bedanke ich mich bei allen weiteren Mitarbeitern der Inneren Medizin II.

Meiner Familie danke ich für ihre große Unterstützung.

## **Eidesstaatliche Erklärung**

Hiermit versichere ich, dass ich die vorliegende Arbeit selbstständig angefertigt und keine anderen als die angegebenen Hilfsmittel und Quellen verwendet habe. Die Arbeit wurde bisher in gleicher oder ähnlicher Weise keiner anderen Prüfungsbehörde vorgelegt und auch nicht veröffentlicht.

---

Ulrike Esslinger

Regensburg, den 30.04.2012