

**Systematische Identifizierung und Charakterisierung
ursächlich genetischer Faktoren seltener urorektaler
und gastrointestinaler Fehlbildungen**

Dissertation

zur

Erlangung des Doktorgrades (Dr. rer. nat.)

der

Mathematisch-Naturwissenschaftlichen Fakultät

der

Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von

Rong Zhang

aus

Taixing, VR China

Bonn 2018

Anfertigung mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der
Rheinischen Friedrich-Wilhelms-Universität Bonn

1. Gutachter: PD Dr. Heiko Reutter
2. Gutachter: Prof. Dr. Michael Pankratz

Tag der Promotion: 04.07.2018

Erscheinungsjahr: 2018

Inhaltverzeichnis

Inhaltverzeichnis	I
Abkürzungsverzeichnis	IV
Abbildungsverzeichnis.....	VII
Tabellenverzeichnis	IX
1 Einleitung.....	1
2 Theoretische Grundlagen	3
2.1 Die Variabilität des menschlichen Genoms	3
2.2 Identifizierung hochpenetranter Mutationen bei monogenen Erkrankungen mittels Next- Generation Sequencing	6
2.3 Genomweite Identifizierung hochpenetranter größerer (> 1kb) Deletionen/Duplikationen (Kopienzahlveränderungen) mittels Array-basierter Karyotypisierung	7
2.4 Identifizierung krankheitsassoziiertes Varianten bei multifaktoriellen Erkrankungen mittels genomweiter Assoziationsstudien	8
2.5 Seltene angeborene Fehlbildungen des Urorektal- und Gastrointestinaltraktes.....	9
2.5.1 Klassifikation von urorektalen und gastrointestinalen Fehlbildungen am Beispiel des BEEK, der ARM und der EA/TEF.....	9
2.5.2 Embryologie von BEEK, ARM und EA/TEF	15
2.5.3 Ätiologie und Epidemiologie von BEEK, ARM und EA/TEF	19
2.6 Aktueller Stand der genetischen Forschung urorektaler und gastrointestinaler Fehlbildungen	22
2.7 Zebrafischmodell zur Untersuchung humaner Fehlbildungen	29
3 Zielsetzung.....	31
4 Material	32
4.1 Patienten- und Kontrollkollektive	32
4.2 Funktionelle Datenbanken	35
5 Methoden.....	38
5.1 Nukleinsäure-Isolation	38
5.2 DNA-Verwaltung.....	38
5.3 Konzentrations- und Qualitätsbestimmung von DNA	39
5.4 Molekularbiologische Methoden	39
5.4.1 Auswahl der Oligonukleotidsequenzen.....	39
5.4.2 Polymerase-Kettenreaktion	40
5.4.3 Agarose-Gelelektrophorese	42
5.4.4 Quantitative PCR (qPCR).....	43
5.4.5 Sanger-Sequenzierung von DNA-Fragmenten.....	45
5.4.6 Genotypisierung mittels Sequenom	48
5.4.7 Genotypisierung mittels Illumina	52

5.4.8 Whole-Exome Sequencing (WES)	54
5.4.9 RNA-Sequenzierung (RNA-Seq)	55
5.5 Bioinformatische Methoden	55
5.5.1 GenomeStudio (GS)	55
5.5.2 Detektion der CNVs aus genomweiten SNP-Arrays mittels QuantiSNP	59
5.5.3 Cartagenia Bench Lab CNV	60
5.5.4 Filterkriterien der CNV-Analyse.....	60
5.5.5 Analyse der Exom-Daten mittels VARBANK	63
5.5.6 dbNSFP zur Interpretation der detektierten Varianten.....	63
5.5.7 Transkriptom-Analyse mittels R	64
5.6 Zellenabstammungsanalyse in Mausembryonen	65
5.7 Whole-mount in-situ-Hybridisierung (WISH)	65
6 Ergebnisse	66
6.1 Genetische Untersuchungen bei Patienten mit Blasenektrophie-Epispadie-Komplex (BEEK) ..	66
6.1.1 Replikationsstudie bei Patienten mit klassischer Blasenektrophie (KBE).....	66
6.1.2 Kandidatengenanalyse in Patienten mit BEEK.....	71
6.1.3 „Lineage Analysis of Celllines“ (Zell-Abstammungsanalysen) in Mausembryonen für <i>ISL1</i> .	76
6.1.4 Whole-mount in-situ Hybridisierung (WISH) in Zebrafisch Larven (zfl).....	77
6.2 Genetische Untersuchungen bei Patienten mit VATER/VACTERL-Assoziation und VATER/VACTERL-ähnlichen Phänotypen.....	79
6.3 Genetische Untersuchungen bei Patienten mit Ösophagusatresien mit und ohne tracheo-ösophagealer Fistel (EA/TEF).....	81
6.3.1 Whole-Exome Sequencing (WES) bei Patienten mit Ösophagusatresien mit und ohne tracheo-ösophagealer Fistel (EA/TEF)	81
6.3.2 Transkriptom-Analyse	81
6.3.3 Kandidatengenanalyse für <i>EEF1D</i>	88
7 Diskussion.....	89
7.1 Blasenektrophie-Epispadie-Komplex (BEEK).....	89
7.1.1 Replikationsstudie bei Patienten mit BEEK und Analyse des <i>ISL1</i> -Locus.....	89
7.1.2 Kandidatenanalyse von <i>SLC20A1</i>	91
7.2 VATER/VACTERL-Assoziation/- ähnliche Phänotypen	93
7.3 Ösophagusatresien mit und ohne tracheo-ösophageale Fistel (EA/TEF)	97
7.3.1 Whole-Exome Sequencing (WES) bei Patienten mit Ösophagusatresien mit und ohne tracheo-ösophagealer Fistel (EA/TEF)	97
7.3.2 Kandidatengenanalyse von <i>EEF1D</i>	99
8 Zusammenfassung.....	100
Referenzen	103

Eigene Publikationen.....	122
Appendix A: Material und Methoden	124
Appendix B: Ergebnisse	140
Danksagung	141

Abkürzungsverzeichnis

A	Adenin
ARM	anorektale Malformation
AUS	Australien
ASPE	enzymatische allelspezifische Primerverlängerung
Array-CGH	Array-based Comparative Genomic Hybridization
BAF	B-Allel Frequenz
BEEK	Blasenektrophie-Epispadie Komplex (BEEK)
Bp	Basenpaar
°C	Grad Celsius
C	Cytosin
CAKUT	Congenital Anomalies of the Kidney and Urinary Tract
CF	Component Feature
Chr	Chromosom
CI	Konfidenzintervall (confidence interval)
CNV	Kopienzahl-Variante (copy number variation)
Ct	Cycle Threshold
DE	Deutschland
dNTP	Desoxyribonukleosidtriphosphat
ddNTP	Didesoxyribonukleosidtriphosphat
DNA	Desoxyribonukleinsäure (deoxyribonucleic acid)
E	Epispadie
EA/TEF	Esophageal atresia with or without tracheo-esophageal fistula
ES	Spanien
EtBr	Ethidiumbromid
EtOH	Ethanol
F	vorwärts (forward)
G	Guanin

Gb	Gigabase
GB	Großbritannien
GWAS	genomweite Assoziationsstudie
HKG	Haushaltsgen
hg	Human Genome
hpf	hours post fertilisation
HWE	Hardy-Weinberg Gleichgewicht (Hardy-Weinberg equilibrium)
IT	Italien
kb	Kilobase
KBE	klassische Blasenektrophie
KE	Kloakenektrophie
KO	knockout
LCR	low copy repeat
LOH	Verlust der Heterozygotie (loss of heterocytosity)
LRR	Log R Ratio
MAF	Minor Allele Frequency
Mb	Megabase
mg	Milligramm
min	Minute
mRNA	Boten-Ribonukleinsäure (messenger ribonucleic acid)
µl	Mikroliter
NGS	Next-Generation Sequencing
nsSNV	non-synonymous single-nucleotide variant
OB-HMM	objektive Bayes Hidden Markov Modell
OD	optische Dichte
PCR	Polymerase-Kettenreaktion (polymerase chain reaction)

qPCR	quantitative Polymerase-Kettenreaktion
R	rückwärts (reverse)
RR	relatives Risiko
RT	Raumtemperatur
SAP	Shrimp alkalische Phosphatase (shrimp alkaline phosphatase)
SBE	Einzelbasenverlängerung (single base extension)
SE	Schweden
sek	Sekunde
smMIPs	Single-molecule Molecular Inversion Probes
ssSNV	splicing-site single-nucleotide variant
SNP	Einzelnukleotid-Polymorphismus (single nucleotide polymorphism)
T	Thymin
TBE	Tris-Borat-EDTA
TDT	Transmission-Disequilibrium Test
TMM	edgeR's Trimmed-Mean of M-values
TOF	Flugzeit (time of flight)
UEP	nicht-verlängerter Primer (unextended primer)
UTR	nicht-translatierter Bereich (untranslated region)
UV	Ultraviolett
WES	Whole-Exome Sequencing
WISH	Whole-mount in-situ-Hybridisierung
WT	Wildtyp
Zfl	Zebrafisch Larven

Abbildungsverzeichnis

Abbildung 1 Männliches (A) und weibliches Neugeborenes (B) mit Epispadie.....	10
Abbildung 2 Männliches (A) und weibliches Neugeborenes (B) mit klassischer Blasenektrophie	11
Abbildung 3 Männliches Neugeborenes mit Kloakenektrophie	12
Abbildung 4 Anorektale Malformationen..	13
Abbildung 5 CFs der VATER/VACTERL-Assoziation.....	14
Abbildung 6 Klassifikation der EA/TEF.....	15
Abbildung 7 Aufreinigungsverfahren mittels Agencourt AMPure XP-Kit	46
Abbildung 8 Verlauf der Aufreinigung des Sequenzierproduktes mittels eines CleanSEQ Kit	48
Abbildung 9 Beispiel für die Darstellung von Chromosom 3 eines Patienten im GenomeStudio	57
Abbildung 10 Beispiel für die Darstellung von Chromosom 13 eines Patienten im GenomeStudio	58
Abbildung 11 Beispiel für die Darstellung von Chromosom 18 eines Patienten im GenomeStudio	58
Abbildung 12 CNV-Filter Schritte von der Genotypisierung bis zur Verifikation	61
Abbildung 13 Genomweite Assoziationsstudie in Patienten mit klassischer Blasenektrophie.....	66
Abbildung 14 Graphische Darstellung der assoziierten Region von <i>ISL1</i>	67
Abbildung 15 Expression von <i>isl1</i> während der embryonalen Entwicklung der Maus.....	72
Abbildung 16 Darstellung des Gens <i>ISL1</i> im UCSC Genome Browser.....	73
Abbildung 17 Zell-Abstammungsanalyse von <i>isl1</i> -exprimierenden Zellen in Mausembryonen	77
Abbildung 18 WISH gegen <i>isl1</i> , <i>slc20a1a</i> und in der Kombination beider Sonden in Wildtyp-zfl bei 48hpf	78
Abbildung 19 Heatmap der Expression von Kandidatengenen mittels der Voom-Methode	87
Abbildung 20 Darstellung der Mikrodeletion bei 17q12 in UCSC Genome Browser.....	94
Abbildung 21 Darstellung von der hier identifizierten Deletion und zwei Duplikationen von Publikationen in der chromosomalen Region 22q11.2	95
Abbildung 22 Darstellung der Überlappungen von 13q31.3del bei Patienten 13 und 13q32.3-q33.1del bei Patienten 14 mit der Deletion in 13q31.1-q34 von Dworschak et al. (2013)	96
Abbildung 23 Darstellung der Überlappungen von 7q35del bei Patientin 7 mit der Deletion in 7q35-q36.3 von Brosens et al. (2016).....	97

Tabellenverzeichnis

Tabelle 1 Krickenbeck-Klassifikation der ARM	13
Tabelle 2 PCR-Reaktionsansatz für unterschiedliche Taq-Polymerasen	41
Tabelle 3 Standard-PCR-Programm TD100	41
Tabelle 4 Standard-PCR-Programm CN	41
Tabelle 5 Parameter in Primer3 zur Auswahl der qPCR Primer.....	43
Tabelle 6 qPCR-Reaktionsansatz mit SYBR Green	44
Tabelle 7 qPCR-Programm mit SYBR Green I (Roche)	44
Tabelle 8 Ansatz der Sequenzierungs-PCR	47
Tabelle 9 Standard-Programm der Sequenzierungs-PCR	47
Tabelle 10 Reaktionsansatz der Multiplex-PCR.....	49
Tabelle 11 Temperaturprogramm einer Multiplex-PCR.....	49
Tabelle 12 Reaktionsansatz des SAP-Verdau.....	50
Tabelle 13 Temperaturprogramm des SAP-Verdau.....	50
Tabelle 14 Reaktionsansatz der SBE-Reaktion	51
Tabelle 15 Temperaturprogramm der SBE-Reaktion	51
Tabelle 16 Transformierung des Thetawertes zur BAF	56
Tabelle 17 Ergebnisse der Assoziationsanalysen und der kombinierten Meta-Analyse der KBE Patienten und Kontrollen sowie der Patient-Eltern-Trios	69
Tabelle 18 Ergebnisse vom TDT in 92 Patient-Eltern-Trios	70
Tabelle 19 Ergebnisse von unseren vorherigen und aktuellen Meta-Analysen für den Marker rs6874700.....	70
Tabelle 20 Darstellung und Klassifizierung der Varianten in SLC20A1 bei Patienten mit BEEK.	75
Tabelle 21 Charakterisierung der chromosomalen Mikroaberrationen in 20 Patienten mit VACTER/VACTERL-Assoziation/-artigen Phänotypen	80
Tabelle 22 Liste der 24 bestätigten de novo Varianten durch WES in 31 Patient-Eltern-Trios	83
Tabelle 23 Charakterisierung der de novo Varianten.....	84
Tabelle 24 Ergebnisse ermittelter Kandidatengene mittels der eBayes-Methode	85

Tabelle 25 Ergebnisse von Kandidatengenomen mittels TREAT mit einem „Fold Change Cutoff“ von $\log_2(1,7)$ 86

Tabelle 26 Darstellung des Grads der Konservierung der identifizierten Varianten mithilfe von UniProt92

1 Einleitung

Urorektale Fehlbildungen repräsentieren Fehlentwicklungen der urorektalen Organe. Anorektale Malformationen (ARM) auf der einen und der Blasenektrophie-Epispadie Komplex (BEEK) auf der anderen Seite repräsentieren die schwersten urorektalen Fehlbildungen. In Deutschland werden jährlich etwa 280 Kinder mit einer ARM oder einem BEEK geboren (etwa 1 auf 3.000 Lebendgeburten) (International Clearinghouse for Birth Defects Monitoring Systems, 1987; Jenetzky, 2007). Die Patienten mit ARM bzw. BEEK haben schwerwiegende funktionelle und kosmetische Beeinträchtigungen. Die Therapie stellt primär eine operative Korrektur der Fehlbildungen dar. Trotz der immensen Verbesserung in der operativen Versorgung gehen diese schweren Formen der urorektalen Fehlbildungen häufig ein Leben lang mit primär und/oder sekundär körperlichen sowie sekundär psychosozialen Folgen einher (Lee et al., 2006). Die Ursachen sind heterogen. Sowohl genetische als auch Umweltfaktoren (z.B. Umweltfaktoren) sind an der Entstehung beteiligt (Reutter et al., 2011; Wijers et al., 2010).

Die Ösophagusatresie mit oder ohne tracheo-ösophageale Fistel (engl. esophageal atresia with or without tracheo-esophageal fistula, EA/TEF) ist die häufigste Fehlbildung des oberen Gastrointestinaltraktes. In Deutschland werden jährlich etwa 230 - 250 Kinder mit einer EA/TEF geboren (etwa 1 in 3.000 bis 3.500 Lebendgeburten) (David and O'Callaghan, 1975; Depaepe et al., 1993; Torfs et al., 1995). Möglicherweise verursachen sowohl Umweltfaktoren als auch genetische Faktoren diese Fehlbildung (Felix et al., 2009). Wie bei den urorektalen Fehlbildungen besteht die Therapie primär in einer operativen Korrektur. Wenn eine primäre Anastomose der Speiseröhre und Resektion der tracheo-ösophagealen Fistel erfolgreich ist, kann es zu einer normalen Entwicklung der Nahrungsaufnahme kommen. Bei schweren respektive langstreckigen Atresien kann es im Verlauf zu schweren Schluckstörungen und gravierenden Ernährungsproblemen in den ersten Lebensjahren kommen (Hölscher et al., 2017).

Für den überwiegenden Anteil der Betroffenen mit einer ARM, einem BEEK oder einer EA/TEF konnten die zugrundeliegenden genetischen Ursachen bislang nicht identifiziert werden. Vorangegangene Studien deuten darauf hin, dass die genetischen Ursachen der hier untersuchten Fehlbildungen sehr heterogen sind. Aufgrund des zumeist sporadischen Auftretens der einzelnen Fehlbildungen kommen neben ursächlichen Neumutationen auch

selten rezessiv vererbte Mutationen sowie ein multifaktorieller Erbgang in Frage. Die entsprechenden Ursachen können daher nur dann erfasst werden, wenn systematisch verschiedene molekulargenetische Untersuchungsmethoden angewendet werden, die den verschiedenen Ursachen tragen. Der technologische Fortschritt ermöglicht es Wissenschaftlern heute die genetischen Ursachen dieser meist sporadisch auftretenden seltenen Fehlbildungen systematisch zu untersuchen.

Einen beträchtlichen Beitrag hierzu kann die Exom-Sequenzierung (engl. Whole-Exome Sequencing, WES) leisten, da mittels WES die genetischen Varianten, die die Proteinsequenzen verändert, identifiziert werden können. Eine andere bedeutende Anwendung ist die genom-weite, chip-basierte Analyse, um systematisch größere Neumutationen zu identifizieren. Auf diese Weise werden ursächliche Deletionen oder Duplikationen genomischer DNA aufgezeigt. Des Weiteren kann der Beitrag genetischer Varianten bei zugrundeliegenden multifaktoriellen Erbgängen durch sog. Assoziationsstudien (genom-weit oder im Einzelnen) ermittelt werden. In der vorliegenden Arbeit wurden beide Methoden systematisch angewandt.

Um die zugrundeliegenden biologischen Mechanismen der identifizierten Kandidatengene im Zusammenhang mit der Entstehung der einzelnen Fehlbildung verstehen zu können und um die Pathogenität identifizierter genetischer Varianten beurteilen zu können, ist eine funktionelle Aufarbeitung erforderlich. Ein Teil dieser Untersuchungen ist ebenfalls Bestandteil der vorliegenden Arbeit.

2 Theoretische Grundlagen

2.1 Die Variabilität des menschlichen Genoms

Das menschliche Genom bezeichnet die Gesamtheit der genetischen Information des Menschen und besteht aus ungefähr 3,2 Milliarden Basenpaaren, die sich auf 24 Chromosomen (22 Autosomen, X- und Y-Gonosomen), sowie die mitochondriale DNA, verteilen (The 1000 Genomes Project Consortium, 2012, 2015; The International HapMap Consortium, 2010). Heutzutage ist das menschliche Genom weitgehend sequenziert und erste Untersuchungen deuteten auf ca. 30.000 bis 40.000 proteinkodierende Gene hin (Lander et al., 2001). Auf Grund von Sequenzlücken bleiben die genaue Größe des Referenzgenoms und die Anzahl von Genen bis heute jedoch ungewiss, außerdem wird die Klassifikation der Gene ständig weiter verfeinert (International Human Genome Sequencing Consortium, 2004). Eine aktuelle Veröffentlichung deutet darauf hin, dass nur knapp 20.000 proteinkodierende Gene im Genom von etwa 3,1Gb vorhanden sind (Ensembl, 2017; GENCODE, 2016). Die proteinkodierende Eigenschaft ist nur eine von vielen Funktionen der DNA, die potentiell mit der Entwicklung bestimmter Phänotypen und/oder Krankheiten verbunden sein können und daher Gegenstand aktueller Forschungen ist. Die ausführlichen Informationen über die genetische Variabilität sind zugänglich durch Annotation der menschlichen Referenzsequenz in Genombrowsern wie dem „Ensembl Genome Browser“ (<http://www.ensembl.org/index.html>) und „UCSC Genome Browser“ (<http://genome.ucsc.edu/>) von der Universität California, Santa Cruz. Zusätzlich sind neben der menschlichen Referenzsequenz auch Referenzgenome von verschiedenen Spezies, wie z.B. der Maus und des Zebrafisches, verfügbar (<http://genome.ucsc.edu/>). Des Weiteren bieten diese und andere Datenbanken Annotationen zu beispielsweise regulatorischen Sequenzelementen, Konservierungsgraden, Phänotypen oder Publikationen.

Die genetische Variabilität ist als die interindividuelle Unterschiedlichkeit in der Sequenz des genetischen Materials, der Desoxyribonukleinsäure (engl. deoxyribonucleic acid, DNA) innerhalb einer Art definiert. Das menschliche Genom ist polymorph, d.h. es gibt viele DNA-Sequenzvarianten in verschiedenen Individuen. Diese Varianten sind ein wichtiger Teil der molekularen Grundlage der genetischen Einzigartigkeit jedes Individuums unserer Spezies. Darüber hinaus ist diese genetische Variabilität das molekulare Substrat von

Evolutionsprozessen. Schließlich verursacht diese Variabilität auch Krankheitsbilder oder eine genetische Prädisposition für häufige oder multifaktorielle Erkrankungen und Merkmale.

Die häufigste DNA-Variante ist der Einzelnukleotid-Polymorphismus (engl. single nucleotide polymorphism, SNP), der durch den Austausch einer einzelnen Base entsteht. Der erste SNP wurde im Jahr 1978 von Forschern aus dem Y.W. Kans Labor in der 3'-Region des β -Globin-Gens identifiziert (Kan and Dozy, 1978). SNPs treten sehr regelmäßig verteilt über das gesamte menschliche Genom auf. Durchschnittlich findet man alle 300 Nukleotide einen SNP, d.h. insgesamt gibt es etwa 10 Millionen SNPs im menschlichen Genom. Viele SNPs haben eine Allelfrequenz von mehr als 5% für das seltenere Allel und sind damit relativ häufig (engl. minor allele frequency, MAF). Im Durchschnitt unterscheiden sich zwei haploide Genome in ungefähr 3.000.000 SNPs (Antonarakis, 2010). Neben den häufigen SNPs liegen noch zahlreiche seltene SNPs (MAF zwischen 1% und 5%) und sehr seltene SNPs (MAF < 1%) im menschlichen Genom vor. Die meisten der häufigen Varianten sind in den Sequenzbereichen zwischen den Genen lokalisiert. Sie können als biologische Marker genutzt werden und Wissenschaftlern helfen, die Gene zu identifizieren, die mit den entsprechenden Krankheiten „assoziiert“ sind. Wenn SNPs innerhalb eines Gens oder in einer regulatorischen Region auftreten, können sie eine direkte Rolle bei der Entstehung von Krankheiten spielen, indem sie die Funktion oder die Synthese des kodierten Proteins beeinflussen. Nach derzeitigem Kenntnisstand haben aber die meisten SNPs keinen Einfluss auf die Gesundheit oder die Entwicklung von Menschen. Weitere Variationen stellen Insertionen/Deletionen (Indels), Inversionen und Kopienzahlvariationen (engl. copy number variations, CNVs) dar. Eine Kopienzahlvariation ist eine Veränderung der Anzahl von Kopien eines bestimmten DNA-Abschnittes (Zugewinn oder Verlust genetischen Materials) innerhalb des menschlichen Genoms. Bisher wurden im menschlichen Genom ca. 30.000 CNVs annotiert (Conrad et al., 2010; Mills et al., 2011). Normalerweise liegen zwei Kopien eines Gens innerhalb eines diploiden Chromosomensatzes vor. Bei der Deletion liegt ein Gen in nur einer Kopienzahl vor (Verlust von genetischem Material). Ein Gen kann aber auch in mehr als in zwei Kopien vorliegen, etwa in drei bis vier Kopien, was als Duplikation bezeichnet wird (Zugewinn des genetischen Materials). Das vollständige Fehlen eines Gens wird als eine homozygote Deletion bezeichnet, die Kopienzahl beträgt demnach null. Die Entdeckung von submikroskopischen Kopienzahlvariationen im humanen Genom, hat unsere Perspektive auf DNA-Strukturvariation und Krankheit drastisch verändert (Stankiewicz and Lupski, 2010). Neben Punktmutationen, die zu „missense

oder nonsense“ Veränderungen des DNA-Codes führen und direkt kausal in Zusammenhang mit spezifischen Krankheitsbildern stehen wie z.B. dem CHARGE-Syndrom, oder der Mukoviszidose, oder die Fanconi Anämie stellen CNVs ebenfalls seltene aber wichtige genetische Ursachen definierter Krankheitsbilder dar wie z.B. dem DiGeorge-Syndrom oder der hereditär motorisch-sensorischen Neuropathie (Jongmans et al., 2006; Redon et al., 2006; Rovelet-Lecrux et al., 2006; Singleton et al., 2003). CNVs können zu verschiedenen molekularen Mechanismen führen wie z.B. einem Gen-Dosiseffekt, der Unterbrechung eines Gens, der Fusion von Genen, oder zu einem Positionseffekt. Aufgrund dieser Vielfalt an molekularen Mechanismen spielen CNVs sowohl bei monogenen als auch multifaktoriellen genetisch komplexen Erkrankungen eine Rolle. Des Weiteren kommt den CNVs eine wesentliche Rolle in der Evolution des menschlichen Genoms zu, in dem es zu Genduplikationen-oder -deletionen sowie zum alternativen Spleißen von Exonen kommt (Zhang et al., 2009).

Die Untersuchung der Verteilung häufiger und seltener genetischer Varianten ermöglicht es Einblicke in Prozesse zu gewinnen, die einerseits ursächlich für die genetische Vielfalt sind, und andererseits kausal für die Ätiologie von Erkrankungen sein können (The 1000 Genomes Project Consortium, 2010, 2012).

2.2 Identifizierung hochpenetranter Mutationen bei monogenen Erkrankungen mittels Next-Generation Sequencing

Obwohl ursächlich monogene Erkrankungen seltener sind als die multifaktoriellen Erkrankungen, beeinträchtigen sie in ihrer Gesamtzahl viele Millionen Menschen weltweit (World Health Organisation, 2017). Eine monogene Erkrankung resultiert aus einem genetischen Defekt in einem einzigen Gen. Der genetische Defekt selbst ist normalerweise, abgesehen von seltenen somatischen Mosaiken, in allen menschlichen Körperzellen zu finden. Sofern das Gen auf den Autosomen liegt, folgt der Erbgang den von Mendel beschriebenen Regeln. Hochpenetrante kausale Varianten sind insgesamt selten, da sie im Zuge von Selektionsprozessen mit einer geringeren Wahrscheinlichkeit an folgende Generationen weitergegeben werden. Seit vielen Jahren war die Sanger-Sequenzierung die vorherrschende Technik, um ursächliche hochpenetrante Mutationen bei monogenen Erkrankungen zu identifizieren. Ein Nachteil dieser Technik ist jedoch, dass man nur ein DNA-Fragment zur selben Zeit und bis zu einer Länge von etwa 1000bp untersuchen kann. Dies hatte zur Konsequenz, dass die Identifikation krankheitsverursachender Mutationen weitaus mühevoller war, als dies heute der Fall ist (Petersen et al., 2017). Der technologische Fortschritt hat das Next-Generation Sequencing (NGS) hervorgebracht. Insbesondere das Whole-Exome Sequencing (WES) hat sich zum „Goldstandard“ entwickelt, um hochpenetrante Mutationen bei monogenen Erkrankungen zu identifizieren (Erzurumluoglu et al., 2015). Das WES ermöglicht die Sequenzierung aller proteinkodierenden Regionen, die ungefähr 2% des menschlichen Genoms umfassen, und die Identifizierung sowohl der exonischen Varianten als auch der Varianten in den Spleißstellen (Petersen et al., 2017). Die standardmäßige Analyse aller Gene kann den Bedarf der zeitaufwändigen Auswahl von Kandidatengenen für die routinemäßige Sequenzierung beseitigen. Es wird geschätzt, dass das Exom ungefähr 85% der Mutationen aufweist, die einen maßgeblichen Einfluss auf krankheitsbezogene Merkmale haben (Majewski et al., 2011). Exonische Mutationen sind ursächlich für die überwiegende Mehrheit monogener Erkrankungen. Dabei machen Missense- und Nonsense-Mutationen allein ca. 60% der krankheitserregenden Mutationen aus (Botstein and Risch, 2003; Kuhlenbäumer et al., 2011).

Aktuelle, öffentlich zugängliche Datensätze der menschlichen DNA-Sequenzvariationen umfassen nur einen kleinen Bruchteil aller sequenzierten Proben: Der Exome Variant Server

(EVS, <http://evs.gs.washington.edu/EVS/>), der als Teil des NHLBI Exome Sequencing Project (ESP) erstellt wurde, beinhaltet Frequenzinformationen von 6.503 Exomen. Das 1000 Genome Project (<http://www.internationalgenome.org/1000-genomes-browsers/>) umfasst derzeit die genetischen Daten des gesamten Genoms von 2504 Individuen (<http://www.internationalgenome.org/data>). Der ExAC Browser (<http://exac.broadinstitute.org/>), der vom Exome Aggregation Consortium (ExAC) generiert wurde umfasst die Exome von 60.706 Individuen unterschiedlicher Ethnizität (Fu et al., 2013; Lek et al., 2016).

Datenbanken für genetische Varianten sind nicht nur wichtig für unser Verständnis von evolutionären Prozessen, sondern stellen auch eine wichtige Ressource für die klinische Interpretation der in Patienten identifizierten Varianten dar (Bamshad et al., 2011; MacArthur et al., 2014). Speziell die Frequenzinformationen potentieller Krankheitsvarianten in den verschiedenen Populationen sind für die Bestimmung von Kausalvarianten bei den jeweiligen Patienten überaus wichtig. Die Effizienz solcher Filtervergleiche hängt dann von der Größe der zur Verfügung stehenden Referenzkohorten ab (Lek et al., 2016).

2.3 Genomweite Identifizierung hochpenetranter größerer (> 1kb) Deletionen/Duplikationen (Kopienzahlveränderungen) mittels Array-basierter Karyotypisierung

Im Vergleich zu den kleineren Mutationsereignissen, die durch das Next-Generation Sequencing oder die Sanger-Sequenzierung identifiziert werden können, können CNVs größere Chromosomenregionen abdecken, die eine Reihe von Genen und regulatorische Regionen betreffen und so Effekte auf den Phänotyp induzieren.

Wie bei den häufigen SNPs nimmt man auch für die im menschlichen Genom häufig vorkommenden CNVs an, dass diese funktionell neutral sind oder in einigen Fällen unsere phänotypische Variabilität ohne Krankheitsrelevanz beeinflussen. Seltene CNVs zeigen im Vergleich zu den häufigen CNVs mit höherer Wahrscheinlichkeit eine Assoziation zu Krankheiten. Um so höher die Penetranz eines CNVs ist, desto wahrscheinlicher besteht ein direkt kausaler Zusammenhang zwischen dem Vorliegen des CNVs und dem Auftreten der Erkrankung (Krepischi et al., 2012). Die wachsende Kenntnis über die funktionellen Auswirkungen von CNVs hat das Verständnis über Unterschiede zwischen einzelnen

menschlichen Genomen und deren genetischer Architektur in Zusammenhang mit multifaktoriellen Erkrankungen revolutioniert.

Um die hochpenetranten CNVs genomweit zu identifizieren, können sowohl Array-CGH-basierte (Array-based Comparative Genomic Hybridization) als auch SNP-basierte Arrays verwendet werden. Dabei können mittels Array-CGH nur CNVs detektiert werden, die durch die fixe Anzahl an Hybridisierungssonden per Array abdeckt sind. Im Gegensatz dazu bietet die Anwendung von hochauflösenden SNP-Arrays die Möglichkeit, genomische Regionen zu betrachten, in denen bisher keine bekannten CNVs detektiert wurden.

Um CNVs, die im Rahmen von wissenschaftlichen Studien oder im Rahmen der diagnostischen Abklärung identifiziert wurden besser beurteilen zu können, stehen öffentliche Datenbanken zur Verfügung: DECIPHER (Database of genomic variation and Phenotype in Humans using Ensembl Resources) und DGV (Database of Genomic Variants). Die genaue Beschreibung von DECIPHER und DGV folgt im Kapitel 4.2. Hierbei hat sich DGV als besonders nützlich erwiesen, um im Rahmen von wissenschaftlichen Studien genomweit identifizierte häufige CNVs herauszufiltern und somit den Datensatz auf seltene und damit eventuell mit der Erkrankung assoziierte CNVs zu komprimieren.

2.4 Identifizierung krankheitsassoziierter Varianten bei multifaktoriellen Erkrankungen mittels genomweiter Assoziationsstudien

Genomweite Assoziationsstudien (engl. genome-wide association study, GWAS) haben sich über die vergangene Dekade hinweg zur wichtigsten Methode für die Erforschung der genetischen Architektur multifaktorieller Erkrankungen entwickelt. In vergleichenden Studien von Patienten und ethnisch angeglichene Kontrollen wird untersucht, ob bestimmte Varianten häufiger in Individuen mit einem bestimmten phänotypischen Merkmal, verglichen mit Individuen ohne dieses Merkmal, auftreten. Die Methode ermöglicht es, potentielle Assoziationen zwischen genetischen Varianten und dem untersuchten Phänotyp statistisch zu erfassen (Burton et al., 2007; Bush and Moore, 2012). Hierbei werden die Frequenzen der auftretenden Allele jeder genotypisierten Variante miteinander verglichen. Eine chromosomale Region, die nach Korrektur für multiples Testen eine statistisch signifikant assoziierte Variante enthält, wird als Suszeptibilitätslocus bezeichnet. Die assoziierte Variante

selbst muss nicht direkt den Phänotyp verursachen, da es sich auch um einen Marker handeln kann, der durch die Haplotypblockstruktur im Kopplungsungleichgewicht mit der tatsächlich ursächlichen genetischen Struktur liegt.

Es ist notwendig, die neu entdeckte Risikovariante in unabhängigen Kohorten von Patienten und Kontrollen zu testen und, falls dies möglich ist, mit alternativen Plattformen zu genotypisieren (McCarthy et al., 2008). Eine Replikationsstudie wird als die zuverlässigste Möglichkeit zur Validierung von GWAS-Befunden betrachtet. Sie hilft sicherzustellen, dass eine Genotyp-Phänotyp-Assoziation, die in einer genomweiten Assoziationsstudie beobachtet wird, eine glaubwürdige Assoziation darstellt und kein zufälliger Befund ist. Ein weiteres mögliches Ziel der Replikationsstudie ist, die Abschätzung der Effektstärke zu verbessern (Kraft et al., 2009). Grundsätzlich sollten bezüglich der untersuchten Population und des Phänotyps klare Kriterien zur Gewährleistung der Vergleichbarkeit angewendet werden, um den Status einer Assoziation nach dem Replikationsversuch zu validieren (Chanock et al., 2007).

2.5 Seltene angeborene Fehlbildungen des Urorektal- und Gastrointestinaltraktes

2.5.1 Klassifikation von urorektalen und gastrointestinalen Fehlbildungen am Beispiel des BEEK, der ARM und der EA/TEF

Blasenekstrophie-Eispadie Komplex (BEEK)

Hinsichtlich des Schweregrades wird der BEEK klinisch in drei verschiedene Subtypen unterteilt: Eispadie (E), klassische Blasenekstrophie (KBE) und Kloakenekstrophie (KE) (Ebert et al., 2009; Ludwig et al., 2009; Reutter et al., 2011).

Die Eispadie (E) bezeichnet die Spaltung der oberen Harnröhre bei beiden Geschlechtern und stellt die leichteste Form des BEEK dar. Bei männlichen Patienten sind glanduläre, penile und proximale Formen dargestellt, wobei nur die proximalen Formen mit einer Inkontinenz vergesellschaftet sind. Bei weiblichen Patientinnen liegt zusätzlich zur Spaltung der Harnröhre noch eine Spaltung der Klitoris vor und es können milde Eispadieformen ohne Inkontinenz vorliegen. Ein Verlust des Blasenhalsschließmuskels und ein „Spaltbecken“ können bei beiden Geschlechtern vorkommen. Bei Männern liegt zusätzlich eine Krümmung des Penis zum Bauch

vor. In der Abbildung 1 werden ein männliches und weibliches Neugeborenes mit Epispadie dargestellt.

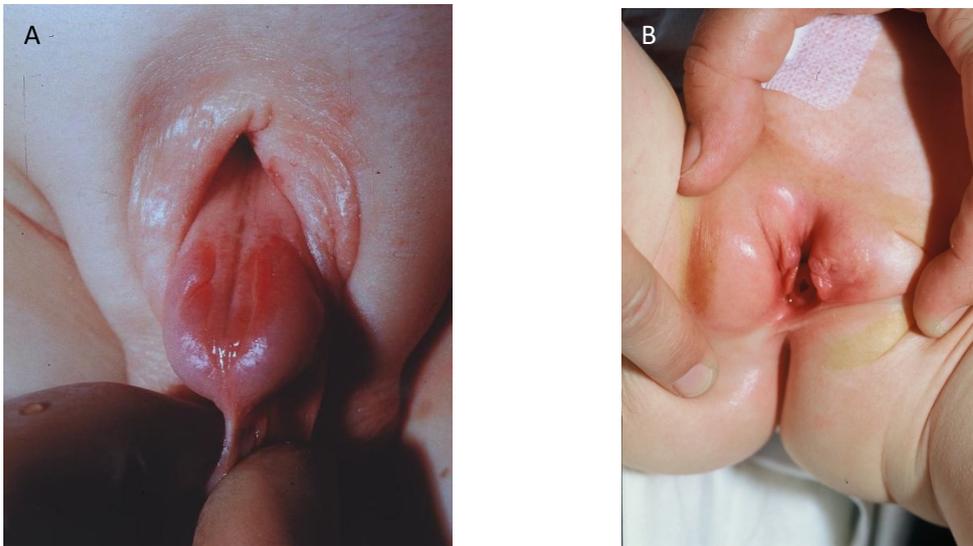


Abbildung 1 Männliches (A) und weibliches Neugeborenes (B) mit Epispadie (Ebert et al., 2009).

Ein klassische Blasenektrophie (KBE) schließt eine komplette Epispadie mit Spaltbecken ein, wobei eine Entwicklungsstörung mit kompletter Spaltbildung des unteren Harntraktes und des äußeren Genitale vorliegt. Im Unterbauch liegt die Blase offen. Der Nabel ist weiter unten zum Oberrand der Blasenplatte verlagert und weist häufig einen Nabelbruch auf. Bei männlichen Patienten liegt die gespaltene Harnröhre als Urethralrinne auf der Doraslseite des Penis an beiden Schwellkörpern. Zur Dorsalseite wird der verkürzte Penis gekrümmt, der bei einer Erektion der unteren Bauchwand anliegt. Bei weiblichen Patientinnen liegen ebenfalls eine gespaltene Harnblase und eine gespaltene Harnröhre vor. Die gespaltene Klitoris ist von der offenliegenden Urethralrinne getrennt. Nach vorne ist der Scheideneingang verlagert und meist etwas verengt. Ein männliches und weibliches Neugeborenes mit klassischer Blasenektrophie werden in Abbildung 2 gezeigt.

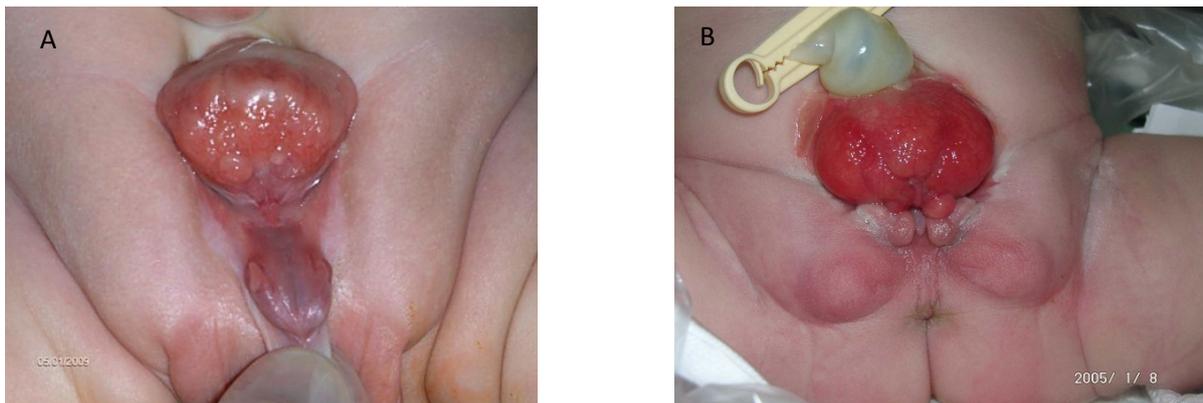


Abbildung 2 Männliches (A) und weibliches Neugeborenes (B) mit klassischer Blasenektrophie (Ebert et al., 2009).

Die Kloakenektrophie (KE) ist die schwerste Form des BEEK, bei der mehrere wichtige Organsysteme betroffen sind. Die klinische Darstellung der Kloakenektrophie ist im Vergleich zu anderen Bauchwanddefekten einzigartig und die KE umfasst neben der Blasenektrophie die Omphalozele und Gastroschisis (Feldkamp et al., 2011). Die KE wird aufgrund der betroffenen Organsysteme auch als „OEIS-Complex“ (Omphalocele, Extrophy, Imperforate anus, Spinal defects) bezeichnet (Carey et al., 1978). Andere Fehlbildungen, die bei der KE häufig auftreten, schließen Nierenfehlbildungen, eine einzelne Nabelarterie und Gliedmaßendefekte ein (Feldkamp et al., 2011; Keppler-Noreuil, 2001; Martínez-Frías et al., 2001; van der Putte et al., 2008). Ösophagusatresien/Tracheo-Ösophagealfistel, Duodenalatresie und Chiari-I-Fehlbildung treten gelegentlich mit auf (Bohring, 2002; McHONEY et al., 2001; Tubbs et al., 2003). Anatomisch sind die Kardinalbefunde der Kloakenektrophie die Ekstrophie der Hemiblasen mit Extroversion des Dickdarms und imperforiertem Anus. Die Hemiblasen flankieren die Öffnungen des Dünndarms und des blindendenden Dickdarms, und enthalten die Öffnungen der Harnleiter bei männlichen Patienten und den Uterovaginalkanal bei weiblichen Patientinnen (van der Putte et al., 2008). In der Abbildung 3 wird ein männliches Neugeborenes mit Kloakenektrophie dargestellt.



Abbildung 3 Männliches Neugeborenes mit Kloakenekstrophie (Ebert et al., 2009).

Anorektale Malformationen (ARM)

ARM sind Entwicklungsdefekte des Enddarms im Stadium der primitiven Kloake dar. Eine Fistel zum Urogenitaltrakt oder eine Verlagerung der Analöffnung tritt häufig bei ARM auf. Wie die anderen Malformationen können ARM entweder isoliert oder in Kombination mit weiteren Fehlbildungen vorkommen. Bei der hier untersuchten VATER/VACTERL-Assoziation (s. nächste Seite) weisen 20 bis 70% der Patienten eine ARM auf (Boocock and Donnai, 1987; Cho et al., 2001; Cuschieri, 2002; Mittal et al., 2004; Stoll et al., 2007).

Traditionell wurden ARM bei atretischem Anus anhand der sogenannten Wingspread-Klassifikation als „hoch“, „intermediär“ und „tief“ in Beziehung zum Musculus levator klassifiziert (Bhatnagar, 2015; Levitt and Peña, 2007; Mittal et al., 2004). Im Jahr 2005, 21 Jahre nach der Wingspread-Klassifikation, wurde eine verbesserte, die sog. Krickenbeck-Klassifikation eingeführt (Tabelle 1) (Holschneider et al., 2005). Die Abbildung 4 gibt die Krickenbeck-Klassifikation anatomisch detailliert wieder.

Tabelle 1 Krickenbeck-Klassifikation der ARM

Hauptgruppe	Seltene/regionale Fehlbildungen
Perineale (kutane) Fistel	Pouch colon
Rekto-urethrale Fistel	Rektumatresie (<1%)/-Stenose
Vestibuläre Fistel	Rekto-vaginale Fistel (<1%)
Bulbäre Fistel	H-Fistel Typ
Prostatische Fistel	Andere
Rekto-vesikale Fistel	
Kloake	
Anorektale Malformation ohne Fistel	
Analstenose	

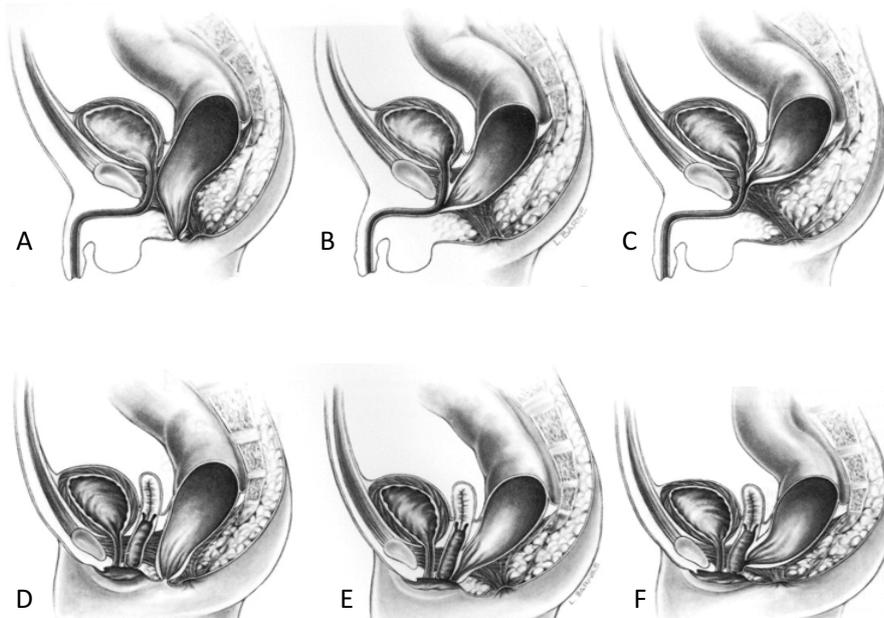


Abbildung 4 Anorektale Malformationen. A: Analatresie mit recto-perinealer Fistel beim Mann; B: Analatresie mit recto-bulbärer Fistel beim Mann; C: Analatresie mit recto-prostatischer Fistel beim Mann; D: Analatresie mit recto-perinealer Fistel bei der Frau; E: Analatresie mit recto-vesikaler Fistel bei der Frau; F: Analatresie mit recto-vaginaler Fistel (eigene Darstellung nach Tobias et al., 2008).

Das Akronym VATER/VACTERL-Assoziation (OMIM #192350) bezieht sich auf das seltene gemeinsame Auftreten folgender phänotypischer Merkmale (engl. component features, CFs): angeborene Defekte der Wirbelsäule (engl. vertebral, V), ARM (Analatresie, A), Fehlbildungen

des Herzens (engl. cardiac, C), tracheo-ösophageale Fistel mit oder ohne Ösophagusatresie (engl. tracheoesophageal fistula with or without esophageal atresia, TE), Anomalien der Nieren (engl. renal, R) und Defekte der (radialen) Gliedmaßen (engl. limb, L) (Abb. 5) (Quan and Smith, 1973). Bei diesen Patienten können zusätzlich weitere kongenitale Fehlbildungen auftreten (Reutter et al., 2014a). Die klinische Diagnose dieser Assoziation erfordert die Anwesenheit von mindestens drei CFs (Quan and Smith, 1973; Hilger et al., 2013; Reutter et al., 2016a; Solomon et al., 2013, 2014). Patienten mit zwei CFs werden als VATER/VACTERL-artige Phänotypen bezeichnet (Reutter et al., 2014a).

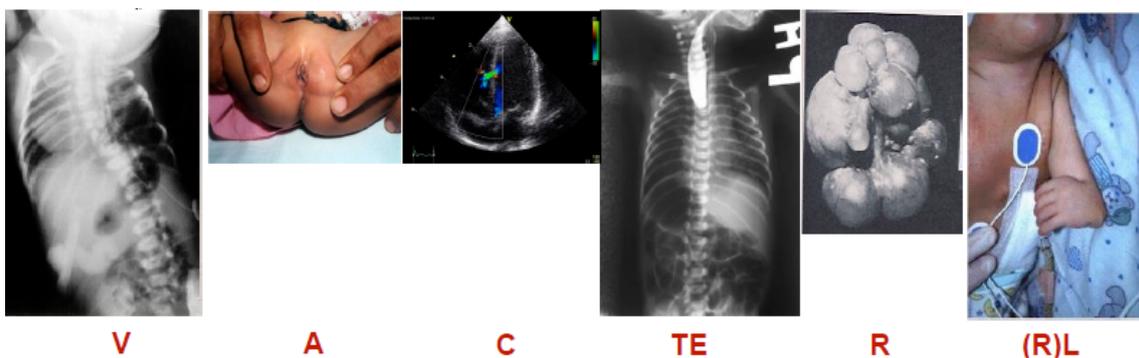


Abbildung 5 CFs der VATER/VACTERL-Assoziation: angeborene Defekte der Wirbelsäule (engl. vertebral, V), ARM (Analatresie, A), Fehlbildungen des Herzens (engl. cardiac, C), tracheo-ösophageale Fistel mit oder ohne Ösophagusatresie (engl. tracheoesophageal fistula with or without esophageal atresia, TE), Anomalien der Nieren (engl. renal, R) und Defekte der (radialen) Gliedmaßen (engl. limb, L) (Bildmaterial POSSUM 5.7.1.).

Ösophagusatresie mit oder ohne tracheo-ösophagealer Fistel (EA/TEF)

Die EA/TEF werden anatomisch durch die Lage der Atresie und die Anwesenheit einer assoziierten Fistel zur Trachea klassifiziert. In dieser Hinsicht wurden fünf verschiedene Typen klinisch beschrieben. Die erste Klassifikation wurde von Vogt im Jahr 1929 veröffentlicht und wurde im Jahr 1953 von Gross modifiziert. Daher werden heutzutage zwei Klassifikationen verwendet. Die häufigste Form der EA/TEF stellt der Typ IIIb nach Vogt dar, der in etwas 85% der Fälle vorkommt (Abb. 6C). Dieser wird gefolgt von dem Typ II nach Vogt mit einer isolierten EA ohne TEF (8%) (Abb. 6A). Die weiteren Subtypen verteilen sich wie folgt: TEF ohne Atresie oder H-Typ TEF (4%, Abb. 6E), EA mit proximaler TEF (3%, Vogt III, Abb. 6B) und EA mit proximaler und distaler TEF (<1%, Vogt IIIa, Abb. 6D). Das Verständnis dieser anatomischen Typen ist notwendig für die chirurgische Rekonstruktion und das neonatale Management nach der Geburt (Pinheiro et al., 2012).

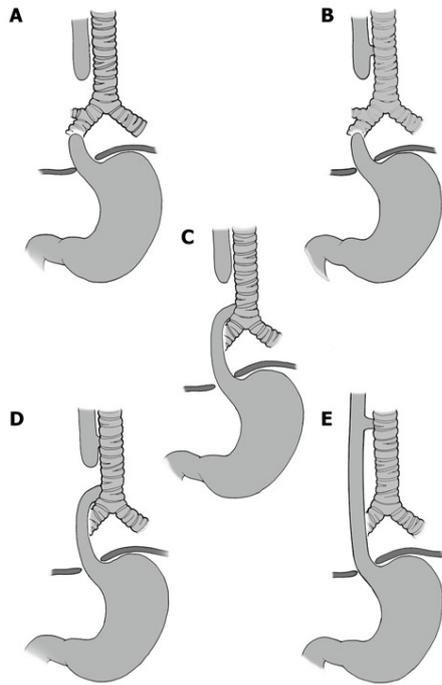


Abbildung 6 Klassifikation der EA/TEF. A: EA ohne TEF; B: proximale TEF mit distaler EA; C: distale TEF mit proximaler EA; D: proximale und distale TEF; E: TEF ohne EA oder „H“-Typ TEF (Pinheiro et al., 2012).

2.5.2 Embryologie von BEEK, ARM und EA/TEF

Blasenektrophie-Epispadie Komplex (BEEK)

Im Jahr 1964 wurde bereits von Muecke berichtet, dass die mechanische Zerstörung oder Vergrößerung der Kloakenmembranen bei Hühnern das Einwandern mesodermaler Zellen entlang der infra-umbilikalen Mittellinie verhindert und damit zur Ekstrophie führt (Muecke, 1964). Tiermodelle zu uro-rektalen Fehlbildungen suggerieren, dass eine Entwicklungsstörung im Bereich der Kloakenmembran die Organisation und Entwicklung des Genitaltuberkels beeinflussen kann und damit zur Entstehung einer Epispadie führen könnte. Dementsprechend wurde auf der Grundlage einer Entwicklungsstudie von erblichen anorektalen Fehlbildungen bei Schweineembryonen abgeleitet, dass eine abnorme Anlage der Kloakenmembran zu verschiedenen angeborenen Fehlbildungen wie z.B. der Hypospadie, Epispadien, KBE, KE, oder der Ausbildung einer Doppelharnröhre führen kann (van der Putte, 1986).

Darüber hinaus haben Thomalla et al. die Kloakenmembran mittels Laser so eingeschnitten, dass ein Herniendefekt in der unteren Bauchwand von Hühnerembryonen (Thomalla et al.,

1985) resultierte, so dass diese bei Geburt einen KE ähnlichen Phänotyp aufwiesen. Diese Beobachtungen unterstütze die Idee von einer vorzeitigen Ruptur der Kloakenmembran bei der Entstehung der KE. Man nimmt an, dass die unterschiedlichen Formen des BEEKs auf eine Ruptur der sehr verletzlichen Kloakenmembran zurückzuführen sind, deren Zeitpunkt und Lokalisation über die Ausdehnung des Defektes entscheidet (Martínez-Frías et al., 2001). Findet die Ruptur statt, ehe sich das Septum urorectale ausgebildet hat, resultiert eine KE mit zwei, den offenliegenden Enddarm flankierenden Hemiblasen. Eine Trennung von Enddarm und Harnblase ist nicht mehr vorhanden, entsprechend einer ekstrophen Kloake (Beetz et al., 1998). Nach Ausbildung des Septum urorectale führt die Ruptur der Kloakenmembran zur KBE; sind lediglich die kaudalen Anteile von der Ruptur betroffen, so entsteht eine Epispadie. Bei hoher Ruptur bildet sich eine so genannte obere Blasenfissur aus, bei welcher lediglich der kraniale Blasenanteil nach außen offen liegt (Beetz et al., 1998).

Ein komplett anderer Mechanismus für die Organogenese der KBE wurde von Beaudoin et al. beschrieben (Beaudoin et al., 2004), die sich dabei auf Beobachtungen der Beckenentwicklung in Kaninchenembryonen beziehen. Sie stellen die zeitlich-räumliche Entwicklung der KBE zu zeitlich-räumlich ablaufenden Anomalien der Beckenknochen in Zusammenhang (Beaudoin et al., 2004). Die Fehlentwicklung soll dabei im Rahmen der sekundären Gastrulation ohne Beteiligung der Kloakenmembran stattfinden. Zusätzlich kann eine KE in Hühnerembryonen durch Zugabe von Nigericin oder Ochratoxin A induziert werden (Vedel-Macrande and Hood, 1986; Wei and Sulik, 1996). Männer und Kluth konnten darüber hinaus eine KE durch die Behandlung von Hühnerembryonen mit Suramin oder Trypanblau in sechs Leghorn-Hühnern erzeugen (Männer and Kluth, 2003). In weiterführenden Untersuchungen an Hühnerembryonen konnten Männer und Kluth zeigen, dass die Exposition der Hühnerembryonen mit Suramin zu einer aneurysmatischen abnorm großen Schwellung der dorsalen Bauchaorta führte, die zu einer abnormen Distension des Beckengürtels und der infra-umbilikalen Bauchwand führt. Dies wiederum könnte zur Ausdünnung der ventralen Bauchwand zwischen dem Umbilikus und der Kloakenmembran führen, welche dann zur Ruptur der Bauchwand führt und im phänotypischen Bild einer kloakalen Ekstrophie mündet (Männer and Kluth, 2005).

Bis vor kurzem gab es nur wenige Beschreibungen über mögliche genetische Ursachen für den BEEK. 2002 konnten Ince et al. (Ince et al. 2002) zeigen, dass *p63* nicht nur eine wichtige Rolle

in der frühen embryonalen Formation des Uro-Rektums darstellt, sondern auch, dass durch komplettes Ausschalten von p63(-/-) im murinen Knockout-Modell in einem Phänotyp resultiert, der in weiten Teilen eine kloakale Ekstrophie beim Menschen widerspiegelt. In einem weiteren murinen Knockout-Modell der Isoform $\Delta Np63(-/-)$ konnten Cheng et al. (2006) das charakteristische Bild einer KBE komplett nachstellen (Cheng et al., 2006).

Anorektale Malformationen (ARM)

Die frühe embryonale Entwicklung des Anorektums, des primitiven urogenitalen Sinus und des kaudalen Neuralrohres steht in engem Zusammenhang und hilft bei der Erklärung der assoziierten Fehlbildungen dieser drei Systeme (Nievalstein et al., 1993, 1994). Zwischen der vierten und sechsten Schwangerschaftswoche treten sowohl der Dottersack oder der primitive Dickdarm, als auch der allantoische oder primitive urogenitale Sinus in die Kloake ein. Das Septum urorectale umkleidet dann (die sog. Tourneux- und Rathke-Falten) die lateralen Kloakenwände, die zum Septum urorectale fusionieren. Gleichzeitig beginnt sich der Embryo als Folge des Längenwachstums des wachsenden Neuralrohres und des mesodermalen Kompartiments zu krümmen. Mit diesen morphologischen Veränderungen wird der Abstand zwischen der Kloakenmembran und der Spitze des Septum urorectale fortschreitend reduziert.

Am Ende der siebten Schwangerschaftswoche befinden sich das Septum urorectale und die Kloakenmembran auf dem gleichen Niveau. Die Kloake ist in einen ventralen (den urogenitalen Sinus) und einen dorsalen Teil (das Rektum und den proximalen Analkanal) unterteilt. Dazwischen entwickelt sich die Spitze des Septum urorectale zum Dammbereich. Zu dieser Zeit bricht die Kloakenmembran durch Apoptose auf und es werden zwei Öffnungen im Perineum gebildet: eine ventrale oder später „urogenitale“ und eine dorsale oder später „anale“. Zunächst durch die Adhäsion der Wände, später durch die Bildung eines epithelialen „Stöpsels“ auf der analen Ebene findet eine sekundäre Okklusion des Anorektalkanal auch zu diesem Zeitpunkt statt (Alamo et al., 2013). Diese sekundäre geschlossene Analöffnung wird durch Apoptose am Ende der achten Schwangerschaftswoche aufgelöst und rekanalisiert.

Aus embryologischer Sicht können ARMs in zwei Hauptgruppen unterteilt werden, je nachdem, wann die Störungen auftreten (Nievalstein et al., 2002, 1998). Diejenigen, die sich als

ektopische Analöffnung oder Fistel manifestieren, sind von der vierten bis siebten Schwangerschaftswoche wahrscheinlich auf eine frühzeitige abnormale Entwicklung des dorsalen Teils der Kloake und der Kloakenmembran zurückzuführen. Ein abnormaler Anus in normaler Position tritt wahrscheinlich als Folge einer späteren fehlerhaften Rekanalisation der sekundär verschlossenen Analöffnung in der siebten und achten Schwangerschaftswoche auf (Alamo et al., 2013; Nievelstein et al., 2002, 1998).

Ösophagusatresie mit oder ohne tracheo-ösophagealer Fistel (EA/TEF)

Das primitive Verdauungrohr (engl. primitive digestive tube, PDT) geht aus dem primitiven Endoderm hervor und differenziert sich anschließend in den Ösophagus und die Trachea. Für die dabei ablaufenden Prozesse existieren derzeit drei primäre Hypothesen (Brunner and van Bokhoven, 2005; El-Gohary et al., 2010; Felix et al., 2009; Ioannides and Copp, 2009; de Jong et al., 2010a; Seo et al., 2010; Shaw-Smith, 2010).

Die Erste geht davon aus, dass die Ausstülpungen des Trachealdivertikels mit der PDT beginnt. Das rasche vorwachsen dieses Divertikels nach kaudal führt zur Trennung von Ösophagus und Trachea. Im Rahmen dieses Entwicklungsprozesses sollen sich EA/TEF durch ein fehlerhaftes Wachstum der Trachea ausbilden (Ioannides and Copp, 2009). Die folgenden zwei Hypothesen gehen davon aus, dass die Trachea ventral und die Speiseröhre dorsal von dem distalen bis zu dem proximalen Ende der PDT durch die Bildung eines mesenchymalen Septums auf der Frontalebene der PDT getrennt wird. Ein Fehler in diesem Prozess könnte dann zur Ausbildung einer EA/TEF führen (Ioannides and Copp, 2009).

Die letzte Hypothese steht im Widerspruch zu den ersten Hypothesen in dem sie davon ausgeht, dass sich der Ösophagus initial als eine unabhängige Struktur bildet, die im Folgenden durch einen bislang nicht bekannten Mechanismus atretisch wird. Demzufolge wächst die TEF von der Trachea auf Höhe der Lungenknospe ein und vereinigt sich mit dem Magen, um die Kontinuität des Gastrointestinaltraktes wieder herzustellen (Ioannides and Copp, 2009).

Obwohl syndromale Formen der EA/TEF selten sind, können dennoch funktionelle Untersuchungen der genetischen Ursachen wertvolle Erkenntnisse bezüglich der abnormen Entwicklungsprozesse, die zu EA/TEF führen, liefern. Auf diesem Weg wurden bereits

zahlreiche Gene und molekulare Wirkmechanismen mit der Entwicklung von EA/TEF in Zusammenhang gebracht (Felix et al., 2009). Gene, die mit der Entstehung von EA/TEF Formen in Verbindung gebracht werden sind z.B. der *FOX*-Gene-Cluster, *GLI2* und *GLI3*, *NKX2.1*, *PCSK5*, *SHH*, *SOX2* und *TBX4* (Pinheiro et al., 2012). Eine gestörte Expression dieser Gene oder apoptotische Prozesse, die durch die genannten Gene reguliert werden, sollen ursächliche Faktoren für die EA/TEF darstellen. Darüber hinaus stellen möglicherweise Umweltfaktoren einen Risikofaktor für die Entwicklung von EA/TEF dar (s.u.).

2.5.3 Ätiologie und Epidemiologie von BEEK, ARM und EA/TEF

Blasenekstrophie-Epispadie-Komplex (BEEK)

Unter Europäern wurde die Geburtenprävalenz über das gesamte BEEK-Spektrum auf 1 in 10.000 Lebendgeburten geschätzt (Reutter et al., 2016a). Obwohl die BEEK-Phänotypen auch als Teil komplexer Fehlbildungssyndrome auftreten, handelt es sich in den meisten Fällen um isolierte Formen (ca. 98,5%) (Boyadjiev et al., 2004; Gambhir et al., 2008; International Clearinghouse for Birth Defects Monitoring Systems, 1987). Caton et al. hatten im Bundesstaat New York die Geburtenprävalenz zuzüglich der Schwangerschaftsabbrüche der einzelnen Subtypen E, KBE und KE betrachtet und fanden, dass die Geburtenprävalenz für E bei etwa 1 in 100.000; für KBE bei etwa 1 in 37.000, und bei KE etwa 1 in 300.000 betrug (Caton et al., 2007). Für die E und die KBE wird jeweils eine klare Knabenwendigkeit beschrieben, die in der Verteilung Männer zu Frauen, zwischen 1,5 zu 1 und 6 zu 1 schwankt (Boyadjiev et al., 2004; ICBDM, 1987; Ives et al., 1980). Weniger einheitlich sind die Berichte zur Geschlechterverteilung für KE, für die neben einer Knabenwendigkeit auch eine gleichmäßige Geschlechterverteilung bis hin zur Mädchenwendigkeit beschrieben wurde (Boyadjiev et al., 2004; Gambhir et al., 2008; Martinez-Frias et al., 2001).

Aufgrund der Seltenheit des Phänotyps weißt man noch zu wenig über mögliche Umweltfaktoren. Gemäß heutigem Wissensstand stellen assistierte Reproduktionsverfahren einen Risikofaktor dar (Wood et al., 2003, 2007; Zwink et al., 2013). Des Weiteren waren mütterliches Rauchen und die medizinische Bestrahlung während des ersten Trimenons mit einem schwereren Phänotyp assoziiert, wohingegen die perikonzeptionelle

Folsäureprophylaxe eine schützende Rolle zu spielen scheint (Reutter et al., 2011). Außerdem ist die Verwendung von Clomifen mit dem Auftreten der KE assoziiert (Reefhuis et al., 2011).

Anorektale Malformationen (ARM)

Anorektale Malformationen (ARM) haben eine Geburtenprävalenz von etwa 1 in 1.500 bis 1 in 5.000 (Zwink et al., 2013). Die Geschlechterverteilung liegt bei etwa 1 zu 1,5 (weiblich zu männlich) (Bischoff et al., 2013; Hamid et al., 2007; Kim et al., 2010).

Die genetischen Ursachen der ARM sind weitestgehend unbekannt. In seltenen Fällen können ARM durch chromosomale Anomalien erklärt werden. Des Weiteren finden sich ARM auch im phänotypischen Spektrum seltener monogener Syndrome wider. Das Townes-Brocks-Syndrom z.B. umfasst Ohren-, Gliedmaßen- Anal-, Nieren- und Herzanomalien und wird durch Mutationen im *SALL1*-Gen verursacht (Kohlhase et al., 1998; Liang et al., 2008). Das Currarino-Syndrom bzw. die Currarino-Trias beschreibt das gemeinsame Auftreten anorektaler und sacro-coccygealer Anomalien sowie präasacraler Raumforderungen, die in den meisten Fällen auf Mutationen im Homeobox-Gen *HLXB9* zurückzuführen sind (Belloni et al., 2000; Crétolle et al., 2008; Hagan et al., 2000). Die meisten ARM-Fälle sind jedoch nicht Teil eines Syndroms, sondern präsentieren sich als isolierte Fehlbildungen.

Die Ätiologie ist heterogen. Neben mikroskopischen und submikroskopischen Chromosomenstörungen und selten monogenen Formen geht man davon aus, dass dem Großteil der ARM eine multifaktorielle Genes zugrunde liegt..

Hinsichtlich der Erforschung nicht-genetischer Ursachen konnten Forscher mögliche Umweltfaktoren identifizieren. Zu diesen assoziierten Umweltfaktoren in der Schwangerschaft zählen mütterlicher Alkoholkonsum, Tabakrauch und Kaffeekonsum, Folsäuremangel die Einnahme von Benzodiazepinen (z.B. Lorazepam) sowie die mütterliche Exposition gegenüber industriellen Reinigungsmitteln (Bonnot et al., 2001; Miller et al., 2009; Myers et al., 2001; Schnitzer et al., 1995; Stoll et al., 1997; Yuan et al., 1995). Außerdem wurden erhöhte Risiken nach Anwendung assistierter Reproduktionsverfahren festgestellt (Källén et al., 2005; Midrio et al., 2006; Reefhuis et al., 2009). Wirklich aussagekräftige Hinweise sind jedoch nur spärlich vorhanden, da die meisten potenziellen Risikofaktoren nur

in einer einzigen Studie gefunden wurden und viele sich bei der Analyse anderer Kohorten nicht replizieren ließen (Wijers et al., 2010).

Ösophagusatresie mit oder ohne tracheo-ösophageale Fistel (EA/TEF)

EA/TEF zeigt eine Geburtenprävalenz von etwa 1 in 3.000 bis 3.500 (David and O'Callaghan, 1975; Depaepe et al., 1993; Torfs et al., 1995). In einer neuen Studie wurde die Geburtenprävalenz von EA/TEF anhand von 18 internationalen Geburtenregistern ermittelt, und es zeigte sich, dass EA/TEF bei ca. 1 von 4.100 Geburten, einschließlich Totgeburten, auftritt. Die Prävalenz fällt zwar über verschiedene geografische Gebiete hinweg unterschiedlich aus, ist aber über die Zeit konsistent und zwischen den Überwachungsprogrammen vergleichbar (Nassar et al., 2012).

Es gibt verschiedene Hinweise, dass Umweltfaktoren in die Entstehung der EA/TEF involviert sind. Eine sechsfach höhere Inzidenz bei Zwillingen, wobei jedoch meist nur ein Zwilling betroffen ist, und das sehr häufig sporadische Auftreten der Erkrankung implizieren die Beteiligung von Umweltfaktoren in der Ätiologie von EA/TEF (Brown et al., 1999; Depaepe et al., 1993; Mastroiacovo et al., 1999; Robert et al., 1993). De Jong et al. (2010) listen mögliche spezifische Umweltfaktoren wie z. B. mütterliche Thiamazol-Exposition (Thyreostatikum), mütterliche Exposition gegenüber Alkohol, Drogen oder infektiösen Erkrankungen, oder eine *in-utero*-Exposition zu Diethylstilbestrol (Di Gianantonio et al., 2001; Felix et al., 2008; de Jong et al., 2010a; Kyyrönen and Hemminki, 1988; Nora et al., 1978). Außerdem wurden fetale Intoxikationen bei Schwangerschaftsdiabetes und die Hypoperfusion in den frühen Phasen der Embryogenese bei Zwillingen mit der Entstehung von EA/TEF assoziiert (Davies and Giménez-Scherer, 2011; Felix et al., 2007; Geneviève et al., 2007). Aktuelle Studien des EUROCAT-Monitoring Systems zur Registrierung angeborener Fehlbildungen (engl. European Surveillance of Congenital Anomalies, EUROCAT) beschreiben, dass ältere Mütter ein deutlich höheres Risiko haben als jüngere Mütter, ein Kind mit EA/TEF zur Welt zu bekommen (Loane et al., 2009).

2.6 Aktueller Stand der genetischen Forschung urorektaler und gastrointestinaler Fehlbildungen

Blasenekstrophie-Eispadie Komplex (BEEK)

Der BEEK stellt eine seltene und schwerwiegende angeborene Fehlbildung dar, die die betroffenen Individuen lebenslang beeinträchtigt. Unter anderem ist die Reproduktionsfähigkeit der Patienten mit BEEK eingeschränkt. Daher wird vermutet, dass *de novo*-Ereignisse eine Ursache für das Auftreten des BEEK darstellen. Die Erfahrung in der medizinischen Genetik zeigen, dass Mutationsereignisse genomische Veränderungen mit unterschiedlichen Größen umfassen, die von kleinen Veränderungen beispielsweise Punktmutationen bis zu großen Veränderungen, mit Zugewinn oder Verlust von mehreren Tausend bis Millionen von Basenpaaren, reichen (Reutter et al., 2016b).

Die Vermutung, dass der Ätiologie des BEEK neben Umweltfaktoren auch genetische Faktoren zugrunde liegen wurde schon früh von verschiedenen Beobachtungen unterstützt. So wurden früh chromosomale Aberrationen bei verschiedenen Patienten mit BEEK beschrieben. Ebenso wurden ca. 30 Familien mit mehreren betroffenen Personen beschrieben. Auch die Beobachtung höherer Konkordanzraten bei monozygoten im Vergleich zu dizygoten Zwillingen weist auf eine genetische Beteiligung hin (Ludwig et al., 2009; Reutter et al., 2003, 2007). Mittels Array-basierter molekularer Karyotypisierung wurden in den letzten Jahren größere strukturelle chromosomale Aberrationen in Patienten mit BEEK identifiziert: Im Jahr 2004 haben Thauvin-Robinet und ihre Kollegen eine relativ große Deletion in 9q34.1 bei einem Säugling mit KE entdeckt (Thauvin-Robinet et al., 2004). Im Jahr 2005 wurde eine erste kleinere Deletion in 3q12.2q13.2 mit einer Größe von 13Mb bei einem Patienten mit KE beobachtet (Kosaki et al., 2005). Eine noch kleinere Deletion mit einer Größe von 2,4Mb wurde bei einem KE-Patienten in der chromosomalen Region 1p36.33 identifiziert (El-Hattab et al., 2010). Außer diesen obengenannten Deletionen ist bislang keine weitere chromosomale Anomalie mit KE assoziiert (Reutter et al., 2016b). Unabhängig davon wurde eine Mikroduplikation in 22q11.2 bei zwei unabhängigen Patienten mit KBE gefunden (Draaken et al., 2010; Lundin et al., 2010). Des Weiteren wurde diese Duplikation in 22q11.21 zusätzlich bei sechs weiteren Patienten mit KBE identifiziert (Draaken et al., 2014; Reutter et al., 2016b). Durch Array-basierte Sequenzierung und gezielte Hochdurchsatz-Re-Sequenzierung konnte gezeigt werden, dass die Bruchpunkte aller dieser 22q11.21 Duplikationen innerhalb

der Low-Copy-Repeats 22A bis 22D lokalisiert sind. Ein Vergleich dieser acht Duplikationen ermöglichte die Eingrenzung auf eine 414 kb-große phänokritische Region, die 12 validierte RefSeq-Gene beherbergt. Die Charakterisierung dieser 12 Kandidatengene mittels In-situ-Hybridisierung (engl. whole-mount in situ hybridisation, WISH) von Mausembryonen am embryonalen Tag 9,5 und 10,5 grenzte diese Kandidaten für die KBE auf *CRKL*, *THAP7* und *LZTR1* ein. Ein weiterer, neunter KBE-Patient, der eine 22q11.21-Duplikation trägt, wurde von Pierquin und Uwineza berichtet (Reutter et al., 2016b). Array-basierte Karyotypisierungen in 279 Patienten mit BEEK deckten weitere CNVs auf, einschließlich einer *de novo* Mikroduplikation mit einer Größe von 0,9Mb in 19p13.12 und einer weiteren Mikroduplikation 22q11.1 mit einer Größe von 1,18Mb (Draaken et al., 2013; von Lowtzow et al., 2016).

Die Befunde aus CNV-Analysen liefern einen Hinweis, dass seltene *de novo* Mutationen zur Ausprägung des BEEK beim Menschen beitragen. Um kleinere dominante *de novo* oder erbliche autosomal-rezessive Mutationen oder kleine CNVs zu identifizieren, erscheint die Verwendung der NGS, insbesondere WES plausibel. Das erste WES wurde von unserer Arbeitsgruppe bei acht Patienten mit KE und ihren Eltern durchgeführt. Dadurch wurden drei dominante *de novo* Mutationen in *PRPF38A*, *PRPF8* und *SLC20A1* entdeckt, die als ursächliche Varianten von den Prognoseprogrammen klassifiziert werden. *Slc20a1* wird in den Harnwegen von Mäusen der embryonalen Theiler-Stufe 22 exprimiert, wobei die Expression regional sehr begrenzt ist. Dies deutet darauf hin, dass *Slc20a1* eine Rolle bei der Entwicklung der Harnwege von Mäusen spielt. Darüber hinaus ergab sich kein Hinweis auf eine Beteiligung von *PRPF38A* oder *PRPF8* in der Ätiologie des BEEK. In dieser WES wurden darüber hinaus vier gemischt heterozygote Varianten in *SPTB*, *WFS1*, *CELSR3* und *SETX* identifiziert. Für *CELSR3* existieren Expressionsdaten in Mäuseembryonen, die eine große Überlappung zu den KE-assoziierten Organen zeigen. Zudem sind Mutationen in *SPTB*, *WFS1* und *SETX* nicht mit Fehlbildungen aus dem KE Spektrum assoziiert sind. Daher verbleiben *SLC20A1* und *CELSR3* als Kandidatengene für die Entwicklung des BEEK. Weitere Studien zu *SLC20A1* und *CELSR3*, einschließlich des Screenings in einer größeren Kohorte, als auch funktionelle Studien in Mäusen und Zebrafischen können dabei helfen, die Beteiligung der Gene an der Entstehung des BEEK zu bestätigen (Reutter et al., 2016b).

Wie oben dargestellt, ist der allgemeine Konsens, dass bei der Mehrheit der BEEK-Patienten die genetischen Grundlagen multifaktoriell sind (Boyadjiev et al., 2004). Um die genetischen Faktoren bei multifaktoriellen Erkrankungen identifizieren zu können, wendet man heutzutage genomweite Assoziationsstudien (genome-wide association study, GWAS) an. Alternativ können Einzelmarkeranalysen für bereits existierende Kandidatengene in Assoziationsstudien durchgeführt werden. Für den BEEK existieren bisher nur wenige Kandidatengen-basierte Assoziationsstudien (Qi et al., 2013; Reutter et al., 2006; Wilkins et al., 2012). Basierend auf ihren Vorarbeiten an Knockout-Mäusen haben Wilkins et al. den Beitrag von *TP63* in der Entstehung de BEEK untersucht und Hinweise dafür gefunden, dass *TP63* an der Pathogenese des BEEK beteiligt ist (Wilkins et al., 2012). Außer *TP63* wurde *PERP* identifiziert, dessen Expression in der Blasen durch WISH nachgewiesen werden konnte (Ihrie et al., 2005). Allerdings haben wir keine Hinweise für eine Beteiligung von *TP63* oder *PERP* durch unsere GWAS gefunden (Draaken et al., 2015). In unserer ersten GWAS befindet sich der signifikant assoziierte SNP innerhalb einer Region mit einer Größe von 4kb neben dem *WNT3*-Promotor, eine Region, die unter Amnioten hoch konserviert ist. Verschiedene Varianten mit unklarer Relevanz in dieser 5'-UTR von *WNT3* waren in einer Kohorte von Patienten mit KE von Vlangos et al. identifiziert worden (Vlangos et al., 2011). Darüber hinaus wurden mehrere potentielle Transkriptionsfaktor-bindende Motive im Promotor von *WNT3* aufgedeckt, von denen sich einige zuvor im Gewebe von menschlichen Neugeborenen mit Blasenektrophie als differentiell exprimiert und notwendig für die Förderung des embryonalen Urorektalseptionsverfahren erwiesen (Kang et al., 1997; Nakamura et al., 2011; Qi et al., 2013). Schließlich wurde gezeigt, dass diese Region regulatorische Elemente enthält, die den Wnt-Signalweg durch p63 regulieren (Ferretti et al., 2011). Diese Befunde unterstützen die Annahme, dass es regulatorische Domänen innerhalb dieser intergenen Region gibt, die in der Lage sind, den Wnt-Singalweg über ein konserviertes *WNT3-WNT9B*-p63-Regulierungsmodul im Rahmen der urorektalen und urogenitalen Entwicklungen zu modulieren.

In unserer zweiten GWAS konnten wir den ersten genomweiten signifikanten Locus für KBE in der chromosomalen Region 5q11.1 identifizieren und erhielten einen starken Hinweis für die Hypothese, dass *ISL1* das verantwortliche Kandidatengen in dieser Region ist (Draaken et al., 2015).

Anorektale Malformationen (ARM) am Beispiel der VATER-VACTERL-Assoziation

Die Analyse von Kandidatengenomen, die auf Array-CGH (engl. comparative genomic hybridization) und SNP-Array basierten Untersuchungen beruhen, sowie das WES wurden durchgeführt, um ursächliche Gene und genetische Risikofaktoren, die zur Entwicklung der Krankheit beitragen, zu identifizieren (Brosens et al., 2013; de Jong et al., 2010b; Ngan et al., 2013; Reutter and Ludwig, 2013; Solomon et al., 2013).

Bisher wurden chromosomale Aberrationen bei mindestens 11 Patienten mit allen Merkmalen der VATER/VACTERL-Assoziation durch zytogenetische und molekulare Analysen aufgedeckt. Im Jahr 1977 haben McNeal et al. eine Deletion in 16q13q15 bei einem Patienten mit VATER/VACTERL-Assoziation entdeckt (McNeal et al., 1977). Des Weiteren wurden Deletionen in den chromosomalen Regionen von 5q11.2, 7q35, 13q, 19q13.3 und 20q13.33 aufgedeckt (Dworschak et al., 2013; de Jong et al., 2010b; Peddibhotla et al., 2013; Solomon et al., 2011; Zen et al., 2010). Außer diesen Deletionen wurden drei Duplikationen in 1q41, 2q37.3 und 8q24.3 von Hilger et al. (2013), eine weitere Duplikation in der chromosomalen Region von 9q von Aynaci et al. (1996), sowie eine Duplikation in 22q11.21 von Schramm et al. (2011) identifiziert. Darüber hinaus wurde das Emanuel-Syndrom, das sogenannte Derivative-22-Syndrom, bei zwei männlichen Patienten beobachtet (Prieto et al., 2007). Außerdem wurden Mosaik eines Ringchromosoms 12 oder Ringchromosom 18, sowie eine partielle Monosomie 16p13.3 und partielle Trisomie 16q22 identifiziert (Cinti et al., 2001; Reutter und Ludwig, 2013; van der Veken et al., 2010; Yamada et al., 2009).

Neben den chromosomalen Anomalien können auch Mutationen in einzelnen Genen zur Entwicklung der Fehlbildungen beitragen. In einer Patientin mit TOF (engl. tetralogy of Fallot), bilateraler Hydronephrose, Hydroureter, ARM und Gliedmaßenfehlbildung, einem VATER/VACTERL-artigen Phänotyp, wurde eine heterozygote *de novo* Deletion von 21 Basenpaaren in *HOXD13*, das für ein Downstream-Target von *SHH* codiert, identifiziert (Garcia-Barceló et al., 2008). Bisher wurde jedoch kein anderer Patient mit der VATER/VACTERL-Assoziation mit krankheitsverursachenden Mutationen in *HOXD13* berichtet (Reutter and Ludwig, 2013). Eine heterozygote *de novo* missense-Mutation in *PTEN* fand sich in einem männlichen Patienten mit VATER/VACTERL-Hydrozephalus-Phänotyp (Reardon et al., 2001). Im Jahr 2014 haben Zeidler et al. eine Analyse des Gens *FGF8* bei 49 Patienten mit VATER/VACTERL-Assoziation und 27 Patienten mit VATER/VACTERL-artigen Phänotypen

durchgeführt. Hierbei wurden zwei heterozygote Mutationen (p.Gly29_Arg34dup und p.Pro26Leu) in *FGF8* jeweils in Patienten mit VATER/VACTERL-Assoziation und VATER/VACTERL-artigem Phänotyp aufgedeckt (Zeidler et al., 2014). Die Duplikation ist noch nicht berichtet, während die Mutation p.Pro26Leu bereits in einem Patienten mit Kallmann-Syndrom, der keine CFs der VATER/VACTERL-Assoziation zeigte, beobachtet wurde (Falardeau et al., 2008). Interessanterweise präsentierten beide Patienten einen bilateralen Kryptorchismus, der ein phänotypisches Merkmal im *FGF8*-assoziierten Kallmann-Syndrom ist. Beide Mutationen in *FGF8* waren vom Vater vererbt. Neben der verspäteten Pubertät bei beiden Vätern und dem einseitigen Kryptorchismus in einem der Väter zeigten diese keine weiteren Auffälligkeiten. Dies deutet darauf hin, dass Mutationen in *FGF8* in den Patienten mit VATER/VACTERL-Assoziation schwere Phänotypen verursachen könnten (Zeidler et al., 2014).

Im Jahr 2010 wurde eine Haploinsuffizienz des Gens *LPP* bei einem Patienten mit VATER/VACTERL-Assoziation festgestellt (Arrington et al., 2010). *LPP* kodiert für ‚LIM-domain, preferred translocation partner in lipoma‘, ein Protein, von dem gezeigt wurde, dass es PEA3 (polyomavirus enhancer activator 3 homolog), einen Transkriptionsfaktor der am *SHH*-Signalweg beteiligt ist, binden kann (Guo et al., 2006). Allerdings zeigten sich keine kausalen Sequenzveränderungen nach der Analyse des *LPP*-Gens bei 70 weiteren Patienten mit VATER/VACTERL-Assoziation (Hernández-García et al., 2012).

Mutationen in den Fanconi-Anämie-Komplementationsgruppen können der VATER/VACTERL-Assoziation ähneln (Holden et al., 2006). Es wurde bei über 200 Patienten mit Fanconi-Anämie beobachtet, dass etwa 5% einen VATER/VACTERL-ähnlichen Phänotyp aufweisen. Obwohl die Fanconi-Anämie eine seltene Ursache für den VATER/VACTERL-Phänotyp zu sein scheint, ist dies eine wichtige Diagnose, da die Erkrankung mit einer hohen Mortalität einhergeht.

Neben der klassischen VATER/VACTERL-Assoziation ist eine weitere assoziierte Form mit Hydrozephalus bekannt. Im Hinblick auf die X-gekoppelte Übertragung wird sie als X-gekoppelter VATER/VACTERL-Hydrozephalus (oder VACTERL-H) bezeichnet, der durch Mutationen im Gen *ZIC3* verursacht werden kann. *ZIC3* ist auch verantwortlich für die X-gekoppelte viszerale Heterotaxie-1 (Chung et al., 2011; Wessels et al., 2010). Im Jahr 2010 berichteten Wessels et al. einen Jungen mit einer Expansion einer Polyalanin-Abfolge in *ZIC3*, die mit dem VATER/VACTERL-Phänotyp ohne Hydrozephalus assoziiert war. Bei einer

weiteren, vier Generationen umfassenden Familie mit mehreren betroffenen Mitgliedern wurde eine Deletion in *ZIC3* nachgewiesen. Ein Familienmitglied mit dieser Deletion zeigte alle Merkmale der VATER/VACTERL-Assoziation (Chung et al., 2011).

Als das erste, autosomal-rezessive kausale Gen für das gesamte klinische Bild der VATER/VACTERL-Assoziation wurde das Gen *TRAP1* in zwei Familien mit isoliertem CAKUT (engl. congenital abnormalities of the kidney and urinary tract) sowie drei Familien mit VATER/VACTERL-Assoziation identifiziert (Saisawat et al., 2014). *Trap1* wird in den Nierenepithelen sich entwickelnder Mäusenieren an E13,5 und in der Niere erwachsener Ratten expimiert. Daher unterstützt die Identifizierung von *TRAP1* als höchstwahrscheinlich verursachendes Gen für CAKUT oder die VATER/VACTERL-Assoziation mit CAKUT die Hypothese, dass variable Expressivitäten der krankheitsverursachenden Gene an der Ätiologie der Krankheit beteiligt sind (Saisawat et al., 2014).

Es gibt nur wenige Tiermodelle bzw. Modelle von Knockout-Mäusen, die einen VATER/VACTERL-artigen Phänotyp aufweisen können. Dazu gehören genetisch manipulierte Mäuse, die Mutationen in den Genen der Shh-Signalkaskade oder eine heterozygote Ethylnitrosoharnstoff-induzierte Mutation in *Psc5* tragen, bilden Fehlbildungen aus dem VATER/VACTERL-Spektrum aus (Kim et al., 2001; Szumska et al., 2008). Aufgrund dieser murinen Phänotypen wird angenommen, dass diese Gene und ihre Signalwege auch beim Menschen eine Rolle in der Ausbildung der VATER/VACTERL-Assoziation spielen (Reutter and Ludwig, 2013).

Ösophagusatresien mit und ohne tracheo-ösophageale Fistel (englisch: esophageal atresia with or without tracheoesophageal fistula, EA/TEF)

Verschiedene Beobachtungen weisen darauf hin, dass genetische Faktoren eine entscheidende Rolle in der Entwicklung der EA/TEF einnehmen. So sind zahlreiche Multiplex-Familien und ein erhöhtes Wiederholungsrisiko für Geschwister und Kinder betroffener Personen beschrieben (Warren et al. 1979; Szendrey et al. 1985; Depaepe et al. 1993; Torfs et al. 1995; Harper 2004). Einen starken genetischen Einfluss legen auch die unterschiedlichen paarweisen Konkordanzraten von monozygoten und dizygoten Zwillingspaaren nahe (Schulz et al. 2012). Neben diesen formal-genetischen Belegen für die Beteiligung von genetischen

Faktoren, die an der Ätiologie von EA/TEF beteiligt sind, wurden auch chromosomale Aberrationen in etwa 6% bis 10% der Individuen mit EA/TEF beschrieben (Depaepe et al., 1993; Shaw-Smith, 2006; Shaw-Smith, 2010; Torfs et al., 1995). Hierzu zählen Aneuploidien wie z.B. das Down-Syndrom, Trisomie 13 und Trisomie 18, sowie Trisomie X (Brosens et al., 2014a, 2014b; Scott, 1993). Neben Aneuploidien werden regelmäßig strukturelle chromosomale Aberrationen (z.B. 13q-, 17q- und 22q11-Deletions-Syndrome) beschrieben (de Jong et al., 2010a). Da bei großen Aneuploidien und strukturellen Aberrationen jeweils viele Gene beteiligt sind, ist es schwierig hieraus, ein einziges kausales Gen zu bestimmen. Im Jahr 2016 wurde von unserer Arbeitsgruppe eine Studie zu Kopienzahlvariationen (CNVs) bei 375 Patienten mit EA/TEF durchgeführt. Hier konnten acht seltene CNVs (Duplikationen bei 2q21.1, 3p26.1, 8p22, 16p13.11 und 21q22.2; Deletionen bei 7q35q36.3, 11p14.3 und 13q12.11) bei sechs Patienten identifiziert werden, die alle *de novo* auftraten. Von diesen Aberrationen wurde nahezu zeitgleich die Deletion bei 7q35q36.3 erneut mit EA/TEF assoziiert (Busa et al., 2016). Zusammengefaßt konnten in unserer Studie *de novo* CNVs in 1,55% der Patienten mit isolierter EA/TEF und in 1,62% der Patienten mit EA/TEF und zusätzlichen angeborenen Anomalien beobachtet werden. Darüber hinaus wurden drei (15q13.3, 16p13.3 und 22q11.2) Suszeptibilitätsloci auf Grundlage ihrer Überlappung mit bekannten EA/TEF-assoziierten CNV-Syndromen identifiziert (Brosens et al., 2016). Zusätzlich zu den identifizierten wahrscheinlich kausalen *de novo* CNVs konnten wir 167 seltene CNVs entdecken. Obwohl diese CNVs wohl nicht direkt krankheitsverursachend sind, könnten sie eine Rolle als Modifikatoren in einem Multi-Hit-Modell oder als „zweiter Hit“ in einem rezessiven Modell spielen (Brosens et al., 2016).

Weitere Hinweise auf die Einbeziehung genetischer Faktoren in die Ätiologie von EA/TEF ergeben sich aus der Beobachtung von über 50 verschiedenen EA/TEF-assoziierten genetischen Syndromen, Assoziationen und Sequenzen (de Jong et al., 2010a; Shaw-Smith, 2006). In einigen dieser EA/TEF-assoziierten genetischen Syndrome wurden die kausal involvierten Gene identifiziert, beispielsweise *N-MYC* im Feingold-Syndrom, *GLI2* im Pallister-Hall-Syndrom, *CDH7* im CHARGE-Syndrom und *SOX2* im AEG-Syndrom (van Bokhoven et al., 2005; Jongmans et al., 2006; Motoyama et al., 1998; Williamson et al., 2006).

Mehrere murine Modelle wurden etabliert, um die Ursachen der EA/TEF zu erklären (Brosens et al., 2014a). Einige Entwicklungsgene sind sicher impliziert, so die Effektoren von Vitamin A

(*Rar α* , *Rar β*) und des Sonic-Hedgehog-Signalwegs (*Shh*, *Gli2*), sowie andere Homeobox-beinhalten Transkriptionsfaktoren und entwicklungsabhängige Transkriptionsregulatoren (*Tbx4*, *Sox2*) (de Jong et al., 2010a). Die Untersuchung der menschlichen Homologen dieser Gene konnte einige wenige humane EA/TEF-Kasuistiken auf genetischer Basis erklären.

Obwohl die zugrundeliegenden chromosomalen Anomalien oder Einzelgenstörungen in einigen syndromalen EA/TEF-Phänotypen identifiziert wurden, bleiben die ursächlichen genetischen Mechanismen in der überwiegenden Mehrheit der betroffenen Individuen derzeit unklar. Der tatsächliche Beitrag von ursächlichen CNVs oder kleineren hochpenetranten *de novo* Mutationen in der Entstehung von EA/TEF ist bis jetzt ungeklärt.

2.7 Zebrafischmodell zur Untersuchung humaner Fehlbildungen

Wissenschaftler haben verschiedene Labortechniken angewandt, um die genetischen Ursachen humaner Erkrankungen zu identifizieren. Hierzu werden üblicherweise humane Zellen oder humanes Gewebe von betroffenen Individuen genutzt. Anders als bei genetischen Erkrankungen, die sich erst postnatal manifestieren, bilden sich die allermeisten humanen Fehlbildungen bereits embryonal aus. Postnatal verändert sich dann das betroffene Gewebe in seiner Form und Struktur nicht mehr. Aber um zu entdecken, ob eine Mutation in einem bestimmten Gen die Syndrome eines Patienten hervorrufen kann, sind die experimentellen Tiermodelle erforderlich. Während Mäuse und Ratten als Tiermodelle zur Untersuchung menschlicher Erkrankungen häufig verwendet werden, simuliert man die menschlichen Fehlbildungen in Knockout-Mäusen oder in Zebrafischen (Dooley and Zon, 2000; Zon, 1999).

Der Zebrafisch (*Danio rerio*) ist ein tropischer Süßwasserfisch. Tatsächlich lassen sich 70% der menschlichen Gene im Zebrafisch wieder finden (Howe et al., 2013). Darüber hinaus hat der Zebrafisch als Wirbeltier die gleichen Hauptorgane und Gewebe wie der Mensch. Viele der Gene und Signalwege, die für die Entwicklung solcher Organe und Gewebe erforderlich sind, sind zwischen dem Menschen und dem Zebrafisch hochkonserviert (Barbazuk et al., 2000; Kroeger and Wingert, 2014). Daher eignet sich der Zebrafisch zur Simulation genetisch bedingter humaner Erkrankungen.

Die heranwachsenden Zebrafische sind klein und können auf kleinem Raum untergebracht werden. Ein weiblicher Zebrafisch ist nach ca. drei Monaten geschlechtsreif und kann bis zu 200-300 Eier produzieren (Hill et al., 2005). Die Eizellen der Zebrafische werden im Wasser gelegt und befruchtet, und dies ermöglicht eine Manipulierung der Embryonen auf vielfältige Weise *in vivo*. Die im Einzell-Stadium befruchteten Eier können leicht mit DNA, RNA oder Morpholino Oligonukleotiden injiziert werden, die die genetische Konstitution dauerhaft modifizieren, um so transgene oder „Knockout-Zebrafischlinien“ zu generieren (Bill et al., 2009). Darüber hinaus sind die Zebrafischembryonen nahezu durchsichtig und durch Zugabe von N-Propylthiouracil (PTU) kann die darauffolgende Pigmentierung temporär verhindert werden. Dies ermöglicht eine Beobachtung des Wachstums des Zebrafischembryos unter einem Mikroskop (Kimmel et al., 1995). Außerdem können die fluoreszenzmarkierten oder gefärbten Gewebe in transgenen Zebrafischembryonen mittels eines Mikroskops visualisiert und beurteilt werden.

3 Zielsetzung

Ziel der vorliegenden Arbeit war es, seltene hochpenetrante Kandidatengene und –regionen für BEEK, ARM und EA/TEF zu identifizieren. Mithilfe von Chirurgen und Humangenetikern, die an dieser Studie beteiligt sind, wurden über 1.000 DNA-/Blut-/Speichelproben von Patienten mit BEEK, ARM und EA/TEF für die genetischen Analysen gesammelt. Der Zweck war, neben der Patientenrekrutierung, eine detaillierte Erfassung der Patientenphänotypen sowie die Errichtung einer zentralen DNA-Biomaterial-Bank.

Die im folgenden dargestellten systematischen Untersuchungen beinhalten zum einen die Replikation einer zuvor mittels GWAS identifizierten genomweit signifikanten Region für die klassische Blasenektrophie, sowie die Identifikation von *de novo* Kopienzahlveränderungen (CNVs). Letztere sind als Verlust oder Zugewinn von genetischem Material gekennzeichnet. Dabei konzentrierten sich die CNV Analysen ausschließlich auf CNVs, die kodierende Gensequenzen beinhalten. Des Weiteren wurden mittels NGS-Technologie besonders schwer betroffene Patienten auf das Vorliegen ursächlicher dominanter *de novo* Mutationen untersucht.

Priorisierte Kandidatengene wurden zum einen mittels Sanger-Sequenzierung in größeren Patientenkohorten weiter untersucht sowie bezüglich ihrer Expression in Mäuseembryonen oder Zebrafischlarven funktionell charakterisiert. Letzteres wurde in Kooperation mit den Arbeitsgruppen von Prof. Benjamin Odermatt (Institut für Anatomie, Universität Bonn, Deutschland), PD Dr. Phillip Grote (Institut für Kardiovaskuläre Regeneration, Universität Frankfurt am Main, Deutschland) und Prof. Gen Yamada (Abteilung für Entwicklungsgenetik, Medizinische Universität Wakayama, Japan) durchgeführt.

4 Material

Appendix A enthält eine Liste aller im Rahmen dieser Arbeit verwendeten

- Geräte
- Chemikalien, Puffer, Lösungen
- Ergänzende Materialien zur Analyse von Nukleinsäuren
- Enzyme
- Zebrafischlinien
- Kommerzielle Systeme (Kits)
- Software und Datenbanken

4.1 Patienten- und Kontrollkollektive

Alle Patienten und Kontrollen waren mit der Teilnahme an den genetischen Studien einverstanden. Von allen Familien liegen uns die schriftlichen Einverständniserklärungen vor. Bei Kindern unter 18 Jahren wurden die entsprechenden Einverständniserklärungen von den Eltern unterschrieben. Alle Studien wurden durch die Etik-Kommission der medizinischen Fakultät der Universitätsklinik Bonn genehmigt.

Netzwerk für kongenitale uro-rektale (congenital uro-rectal) Fehlbildungen (CURE-Net)

Ziel des Netzwerks für kongenitale uro-rektale Malformationen (CURE-Net) ist die systematische Untersuchung der molekularen Ursachen und die Evaluierung des postoperativen Behandlungserfolges und der unterschiedlichen Formen der Nachsorge durch multizentrische, klinische und psychosoziale Forschung mit Hilfe von standardisierten Untersuchungserfahren. Seit 2009 arbeiten die klinisch tätigen Ärzte und Grundlagenwissenschaftler aus verschiedenen Fachbereichen deutschlandweit mit den beiden Patientenorganisationen „Selbsthilfegruppe Blasenektrophie/Epispadie e.V.“ (SHG BE e.V.) und „Selbsthilfeorganisation für Menschen mit angeborenen anorektalen Fehlbildungen“ (SoMA e.V.) zusammen. Die BEEK- und ARM-Patienten aus Deutschland wurden durch CURE-Net gesammelt. Für die Analysen in dieser Arbeit wurden von jedem

Patienten Blut- oder Speichelproben gesammelt. Aus den gesammelten Blut- oder Speichelproben wurde dann die DNA extrahiert und asserviert.

Das GREAT Konsortium (Genetic Risk of Esophageal ATresia)

Das GREAT Konsortium wurde von PD Dr. Heiko Reutter und PD Dr. Johannes Schumacher (Insitut für Humangenetik, Bonn) gegründet, um die genetischen Risikofaktoren bei Patienten mit Ösophagusatresien systematisch zu identifizieren. Mittels des GREAT Konsortiums wurden die in dieser Arbeit analysierten Patienten und ihre Eltern in Bonn rekrutiert. Zur Analyse der genetischen Veränderungen wurden von jedem Patienten und den Eltern Blut-, Speichel- oder DNA-Proben gesammelt.

Amerikanisches Kollektiv

Durch die Kooperation mit einem amerikanischen Kollegen aus Nordamerika wurden uns DNA-Proben von 110 Eltern-Trios mit KBE bereitgestellt. Die Personen dieser 110 Trios waren Nordamerikaner mit europäischen Vorfahren.

Australisches Kollektiv

Durch die Kooperation mit Prof. Wei Cheng (Department of Pediatric Surgery, Capital Institute of Pediatrics, Beijing, People's Republic of China; Department of Paediatrics and Department of Surgery, Southern Medical School, Faculty of Medicine, Nursing and Health Sciences, Monash University, Clayton, VIC, Australia; Department of Surgery, Beijing United Family Hospital, Beijing, People's Republic of China) haben wir Zugang zu DNA-Proben von 66 KBE-Patienten europäischer Abstammung.

Britisches Kollektiv

Durch die Kooperation mit Prof. William Newman (Manchester Centre for Genomic Medicine, University of Manchester, Manchester, England) hat unsere Arbeitsgruppe Zugang zu DNA-Proben von 40 britischen KBE-Patienten und 80 gesunden Kontrollpersonen.

Italienisches Kollektiv

Durch die Kooperation mit Prof. Alfredo Brusco (Department of Medical Sciences and Medical Genetics Unit, Città della Salute e della Scienza University Hospital, University of Torino, Torino, Italy) hat unsere Arbeitsgruppe Zugang zu DNA-Proben von 40 italienischen KBE-Patienten und 85 gesunden Kontrollpersonen.

Schwedisches Kollektiv

Durch die Kooperation mit Prof. Agneta Nordenskjöld (Department of Women's and Children's Health and Center for Molecular Medicine, Karolinska Institutet, Stockholm, Sweden) hat unsere Arbeitsgruppe Zugang zu den Genotypen von 116 Schwedischen KBE-Patienten und 177 gesunden Kontrollpersonen, die mittels Taqman-Assays und Sanger-Sequenzierung gewonnen wurden.

Spanisches Kollektiv

Zwischen 2004 und 2006 wurden 35 spanische KBE-Patienten durch PD Dr. Heiko Reutter mit Hilfe der spanischen Selbsthilfegruppe gesammelt (www.extrofia.info/asafex). Mit Hilfe von ESIGEM network (www.esigem.org) haben wir Zugang zu genomweiten Genotypen von 871 spanischen Kontrollen.

Heinz Nixdorf Recall (HNR) Kollektiv

Das HNR-Kollektiv wurde im Ruhrgebiet für die Untersuchung von Herz-Kreislauf-Erkrankungen rekrutiert (Schmermund et al., 2002). Auf dem BeadChip HumanOmniExpressExome-8 v1.2 wurden 1.307 HNR-Kontrollen genotypisiert und QuantiSNP-Daten erhoben. Die im HNR-Kontrollkollektiv erhobenen CNVs wurden für den Vergleich der in unserer Patientenstudie erhobenen CNVs genutzt. CNVs die zwischen Kontrollpersonen und Patienten überlappten wurden ausgeschlossen.

4.2 Funktionelle Datenbanken

In dieser Arbeit wurden verschiedene funktionelle Datenbanken für die Analyse der Kandidatengene und Kandidatenregionen verwendet. Die verschiedenen Datenbanken sind in der Folge im Einzelnen aufgeführt.

Database of Chromosomal Imbalance and Phenotype in Humans using Ensembl Resources (DECIPHER)

DECIPHER (<https://decipher.sanger.ac.uk/>) ist eine interaktive, webbasierte Datenbank, die für die Interpretation der detektierten submikroskopischen chromosomalen Aberrationen geschaffen wurde (Firth et al., 2009). DECIPHER katalogisiert die häufigen Kopienzahlveränderungen in den normalen Populationen und ermöglicht damit den Ausschluss solcher benigner Kopienzahlveränderungen und damit die Priorisierung vermeintlich neuer und potenziell pathogener Veränderungen. Bekannte Gene innerhalb einer Aberration, sowie die phänotypischen Daten der Patienten sind im DECIPHER Patientenbericht aufgeführt, die Gene mit einer anerkannten klinischen Bedeutung sind hervorgehoben und priorisiert (Firth et al., 2009). DECIPHER ist die größte CNV-Datenbank mit phänotypischen Daten. Bisher sind in DECIPHER 27.977 CNVs und 3.370 Sequenzvarianten in 24.323 Patienten unter Beteiligung von mehr als 250 Zentren erfasst (Stand 23.07.2017). Daher wurde DECIPHER in dieser Arbeit verwendet, um die verifizierten CNVs weiter zu analysieren und zu interpretieren.

Database of Genomic Variants (DGV)

Das Ziel des DGV-Projektes (<http://dgv.tcag.ca/dgv/app/about?ref=GRCh37/hg19>) ist eine umfassende Zusammenfassung der strukturellen Variationen im menschlichen Genom zu erfassen. Im DGV sind die genomischen Veränderungen, einschließlich der Segmente der DNA, die größer als 50bp sind, als strukturelle Varianten definiert. Die DGV umfasst nur die strukturellen Variationen, die in gesunden Kontrollproben identifiziert wurden. Daher wurde die DGV in dieser Arbeit dazu eingesetzt, die identifizierten CNVs auszuschließen, die im DGV als benigne klassifiziert wurden.

ExAC (Exome Aggregation Consortium) Browser

Der ExAC Browser (<http://exac.broadinstitute.org/>) ist vom „Exome Aggregation Consortium (ExAC)“ entwickelt worden, um die Exom-Sequenzierungsdaten aus einer Vielzahl von umfangreichen Sequenzierungsprojekten zu sammeln und zu harmonisieren, sowie die Daten für Wissenschaftler weltweit zusammenzufassen. Der auf diesem Browser bereitgestellte Datensatz umfasst die Exom-Sequenzierungsdaten von 60.706 Individuen (Lek et al., 2016). In den hier beschriebenen Untersuchungen wurden die detektierten Varianten mit den Varianten im ExAC Browser hinsichtlich ihrer MAF (engl. minor allele frequency) verglichen, um die seltenen Varianten auszufiltern.

UCSC (University of California Santa Cruz) Genome Browser

Im UCSC Genome Browser (<https://genome.ucsc.edu/index.html>) sind verschiedene genomische Sequenzen und Annotationen für Wirbeltiere erfasst (Kent et al., 2002). Mittels des UCSC Genome Browsers werden die Visualisierung, der Vergleich und die Analyse der CNVs und der Varianten ermöglicht.

Mouse Genome Informatics (MGI) Datenbank

Die MGI Datenbank ist eine internationale Datenbank für die Labor-Mäuse und bietet ein integriertes und kuratiertes System, welches einen aktuellen umfassenden Überblick über Mausgene, genetische Marker und genomische Merkmale sowie die Assoziation dieser Merkmale mit Sequenzen, Phänotypen, funktionellen und vergleichenden Informationen, sowie ihre Beziehungen zu menschlichen Krankheiten anbietet (Blake et al., 2017). In dieser

Arbeit wurden die durch WES und Sanger-Sequenzierung identifizierten Gene mittels der MGI Datenbank im Hinblick auf ihre Expression in Mäuseembryonen verglichen.

The Zebrafish Information Network (ZFIN)

Das ZFIN (<http://zfin.org>) ist die zentrale Ressource für die genetischen, genomischen und phänotypischen Daten sowie Entwicklungsinformationen von Zebrafischen. Mittels ZFIN wurden in dieser Arbeit Morpholinos für einen Knock-Down der Genexpression in Zebrafischen ausgewählt und die durch WES und Sanger-Sequenzierung identifizierten Gene im Hinblick auf ihre Expression verglichen.

5 Methoden

5.1 Nukleinsäure-Isolation

In dieser Arbeit wurde die hochmolekulare DNA aus den kernhaltigen Leukozyten des Probandenblutes bzw. aus den im Speichel befindlichen Leukozyten und die RNA aus Gewebe extrahiert.

Mittels EDTA-haltigen Monovetten (SARSTEDT AG & Co.) wurde venöses Vollblut aufgezogen und bis zur Isolierung der DNA wurden die EDTA-Blutproben bei -20°C aufbewahrt. Mittels Oragen-Kit OG-250 und OG-500 (DNA Genotek, OraSure Technologies) wurden erst die Speichelproben gesammelt und nach der Inkubation bei 50°C bis zur abschließenden Extraktion der DNA bei RT gelagert. Die Isolierung von DNA aus Blut wurde nach Herstellervorgaben größtenteils mit dem Chemagic Magnetic Separation Module I v20 (PerkinElmer, Baesweiler) und dem „QIAamp DNA Blood Maxi Kit“ durchgeführt. Die DNA-Präparation aus Speichel erfolgte nach Herstellervorgaben mit dem entsprechenden Kit.

Embryonen aus schwangeren Weibchen des C57Bl6J-Mausstammes wurden an den embryonalen Tagen E8.5 (entsprechendes Entwicklungs-Theiler-Stadium 13, TS13) und E12.5 (TS21) entnommen. Aus den TS13-Embryonen wurde der Pharynx Beutel, der Endoderm- und angrenzendes Mesodermgewebe enthielt, chirurgisch isoliert und in QIAzol[®] transferiert. Aus den TS21-Embryonen wurde die ausgeprägte Struktur des Ösophagus chirurgisch isoliert und in QIAzol[®] transferiert. Die gesamte RNA wurde aus diesen Geweben mit dem RNeasy Plus Mini Kit (Qiagen) gemäß den Herstellersprotokollen isoliert.

5.2 DNA-Verwaltung

Von jeder DNA-Probe wurden nach der Isolierung zwei Verdünnungen hergestellt. Aus der DNA-Stocklösung wurde manuell bis zur Konzentration von $100\text{ng}/\mu\text{l}$ in TE^{-4} verdünnt. Die $100\text{ng}/\mu\text{l}$ Verdünnungen wurden bei -20°C aufbewahrt. Die zweite Verdünnung wurde mit Hilfe vom Pipettierroboter in H_2O mit einer Endkonzentration von $20\text{ng}/\mu\text{l}$ hergestellt und bei 4°C gelagert. Die $20\text{ng}/\mu\text{l}$ und $100\text{ng}/\mu\text{l}$ Verdünnungen dienten als Ausgangsmaterial für die in dieser Arbeit durchgeführten molekulargenetischen Analysen. Alle Verdünnungen wurden mit dem 2D-Barcode-System von Abgene und FluidX verwaltet.

5.3 Konzentrations- und Qualitätsbestimmung von DNA

Qualität und Quantität der DNA wurden mittels des NanoDrop ND-1000 bzw. des ND-8000 (Peqlab Biotechnologie GmbH) Spektrophotometers unter Verwendung von 1,5µl DNA-Lösung pro Messung bestimmt. Die Konzentrationsmessung basiert auf der Absorptionseigenschaft von DNA, deren Absorptionsmaximum bei einer Wellenlänge von 260nm liegt. Weitere Moleküle wie Proteine und Salze, die nach der Isolation der DNA in der Lösung verbleiben, weisen ein Absorptionsmaximum von 280nm bzw. 230nm auf. Anhand des Lambert-Beerschen-Gesetzes wird die Konzentration der DNA aus der optischen Dichte (OD) bei 260nm bestimmt. Das OD-Verhältnis bei 260nm und 280nm (OD 260/280) gibt einen Hinweis über die Reinheit der DNA. Eine reine DNA-Probe zeigt ein Quotient von 1,8 an. Bei Bedarf wurde zusätzlich eine Degradierungsanalyse mittels Gelelektrophorese in einem 0,8%igen Agarosegel (0,64g Agarose; 80ml 1XTBE; 6µl EtBr) durchgeführt. Nach 45 Minuten bei 100V (40mA) sollte eine definierte Bande der 100ng eingesetzten hochmolekularen DNA auf dem Agarosegel zu sehen sein. Eine degradierte DNA bildet hingegen eine Schmier auf dem Gel aus.

5.4 Molekularbiologische Methoden

5.4.1 Auswahl der Oligonukleotidsequenzen

Die Auswahl der Oligonukleotidsequenzen ist das sogenannte Primer-Design. Die Oligonukleotide (Primer) sind für eine erfolgreiche Durchführung der Polymerase-Kettenreaktion (PCR) bzw. Sequenzierung notwendig und sie müssen genau komplementär zur Ziel-DNA-Sequenz sein. Normalerweise sollten die folgenden Regeln beim Design der Primer Anwendung finden:

- Aufgrund einer unspezifischen Anlagerung des Primers sollte eine Folge von drei oder mehr G- oder C-Basen am 3'-terminalen Ende des Primersequenz vermieden werden.
- Damit sich keine Primerdimere ausbilden können, sollten die Sequenzen an den 3'-terminalen Enden weder inter- noch intramolekular komplementär sein.

- Eine Vermeidung eines Thymins am 3'-terminalen Ende des Primermoleküls ist notwendig, da im Vergleich zu anderen Nukleotiden ein Thymin an dieser Position zu Falschpaarungen führen könnte.
- Die Länge eines Primers sollte zwischen 18 und 25 Nukleotiden sein und die Differenz zwischen den Annealing-Temperaturen des Primerpaares sollte nicht 2°C übersteigen.
- Im Primer sollten keine häufigen SNPs liegen. Nach Möglichkeit sollte auch kein seltener SNP enthalten sein.
- Die Primersequenz sollte spezifisch komplementär zu der gezielten DNA-Sequenz gebaut werden und nur einmal im ganzen Genom vorkommen.
- Alle vier Basen sollten in den Primer, wenn möglich, gleich häufig vertreten sein.

5.4.2 Polymerase-Kettenreaktion

Die Polymerase-Kettenreaktion (engl. polymerase chain reaction, PCR) ist ein molekularbiologisches Verfahren, um spezifisch DNA-Segmente *in vitro* zu vermehren. Dieses Verfahren besteht aus drei sich 30 bis 35-mal wiederholenden Zyklen, nämlich das Aufschmelzen der DNA, die Anlagerung der Primer an ihre komplementäre Sequenz und die Elongation mit Hilfe von einer DNA-Polymerase. Voraussetzung für eine spezifische Amplifikation ist die Kenntnis der Zielsequenz. Die Primer, die für die Zielsequenz spezifisch sind, wurden mittels der „Primer3 Software“ (<http://primer3.ut.ee/>) generiert. Die in dieser Arbeit verwendeten Primersequenzen sind in Appendix A2 aufgelistet. Die genomische DNA ist als Ausgangsmaterial für die PCR in dieser Arbeit verwendet worden.

In Tabelle 2 ist ein Standard-Ansatz für die PCR dargestellt. Alle Reaktionen wurden in dünnwandigen, sterilen Reaktionsgefäßen und in einem Thermozykler nach dem in den Tabellen 3 und 4 gelisteten Reaktionsprotokollen durchgeführt. Es wurden zwei unterschiedliche Taq-Polymerasen, je nach Fragestellung und Eigenschaft des PCR-Produktes, verwendet. Die Überprüfung des Erfolgs und der Effizienz der Amplifikation wurde anhand der Gelelektrophorese durchgeführt.

Tabelle 2 PCR-Reaktionsansatz für unterschiedliche Taq-Polymerasen

	VWR-Taq DNA-Polymerase (VWR)	HotStar Taq DNA-Polymerase (Qiagen)
DNA (20-100ng)	1µl	1µl
Primer F (10pmol)	1µl	1µl
Primer R (10pmol)	1µl	1µl
dNTP-Mix	0,5µl	0,5µl
10xPuffer	2,5µl	2,5µl (15mM MgCl ₂)
Taq	0,2µl	0,2µl
DMSO (optimal)	1,25µl	1,25µl
Destilliertes Wasser	Ad 25µl	Ad 25µl

Tabelle 3 Standard-PCR-Programm TD100

Reaktionsphase	Temperatur	Zeit
initiale Denaturierung	95°C	5 min
Amplifikation	95°C	30 sek
	63-55°C *	30 sek
	72°C	1 min/kb
finale Elongation	72°C	10 min
Abkühlung	12°C	unendlich

* Initiale Annealing-Temperatur bei 63°C für drei Zyklen (Amplifikation), anschließend jeweils alle drei Zyklen ein Grad kälter, 15 Zyklen am Ende bei der Annealing-Temperatur von 55°C.

Tabelle 4 Standard-PCR-Programm CN

Reaktionsphase	Temperatur	Zeit	Zyklen
initiale Denaturierung	95°C	5 min	1
Amplifikation	95°C	30 sek	35
	primerspezifisch	30 sek	
	72°C	1 min/kb	
finale Elongation	72°C	10 min	1
Abkühlung	12°C	Unendlich	1

5.4.3 Agarose-Gelelektrophorese

Die Agarose-Gelelektrophorese ist ein Elektrophorese-Verfahren, bei dem Agarosegel als Trägermedium die Gemische von DNA-Fragmenten nach ihrer Länge und Ladung auftrennt. Die DNA-Moleküle wandern aufgrund ihrer negativen Eigenladung in polymerisierter Agarose durch das Anlegen einer Spannung in Richtung Anode. Die Wanderungsgeschwindigkeit der DNA-Moleküle im Trägermedium variiert aufgrund der Größe der Fragmente und der Stärke der angelegten Spannung sowie der Konzentration des Agarosegels. Je nach Größe der Fragmente wird das Agarosegel mit entsprechender Konzentration eingesetzt, d.h., kleinere Fragmente können sich genauer durch ein Gel mit erhöhter Agarose-Konzentration auftrennen lassen.

In der vorliegenden Arbeit wurden Agarosegele mit einer Konzentration von 0,8% bis 2% hergestellt. Der fluoreszierende Farbstoff Ethidiumbromid (EtBr) wurde der gelösten Agarose zugesetzt, und EtBr interkaliert in die DNA-Doppelhelix. Anhand des UV-Lichts werden die Banden sichtbar. Eine DNA-Leiter wurde gleichzeitig neben den zu analysierenden Fragmenten aufgetragen, um die Länge der DNA-Fragmente zu bestimmen.

Zur Herstellung eines Gels wurde die entsprechende Menge Agarose in 80ml (240ml) 1X TBE Puffer zugegeben und durch Aufkochen im Mikrowellenherd gelöst. Anschließend wurde es bis auf ca. 40 bis 50°C abgekühlt und 6µl (18µl) EtBr-Lösung zugesetzt. Die flüssige Agarose-Lösung wurde in einen geeigneten Gelträger gegossen. Wenn das Gel polymerisiert war, wurde es in eine geeignete, mit 1X TBE gefüllte Gelelektrophorese Kammer (Bio-Rad) gelegt.

Die aufzutrennenden DNA-Fragmente wurden mit 1X Ladepuffer versetzt und anschließend auf das Gel aufgetragen. Die Elektrophorese wurde bei 80 bis 130 Volt (je nach der Konzentration des Agarosegels) für 30 Minuten durchgeführt. Unter UV-Licht wurden die Fragmentgröße und die DNA-Konzentration im Vergleich zu einer gesondert in eine Geltasche aufgetragenen DNA-Leiter abgeschätzt. Mittels des Gel Doc XR-Systems (Bio-Rad) wurde das Gel fotodokumentiert.

5.4.4 Quantitative PCR (qPCR)

Zur Validierung der identifizierten Kopienzahlveränderungen (CNVs) wurde die quantitative PCR (qPCR) verwendet. Der prinzipielle Ablauf der qPCR ähnelt der einer PCR, jedoch erfolgt hier zusätzlich eine Zugabe von SYBR Green I Master (Roche). SYBR Green I ist ein asymmetrischer Farbstoff, der in doppelsträngige DNA interkaliert. Die detektierbare Fluoreszenzintensität wurde durch diese Interaktion der DNA mit SYBR Green I nach jedem Amplifikationszyklus erhöht und auch gemessen.

Die Reaktionen wurden in einem LightCycler 480 System (Roche) durchgeführt und analysiert. Die vorhandene Fluoreszenz wird nach jedem Zyklus gemessen und so wird der Ct-Wert (Cycle Threshold) gebildet. Der Cycle Threshold ist der Schnittpunkt zwischen einer Amplifikationskurve und einer Schwellenlinie und beschreibt den Anfang des exponentiellen Wachstums einer Amplifikationskurve nach Abzug der Hintergrundfluoreszenz. Anhand eines Schwellenwertes, der für alle Proben gleich ist, liegt der Ct-Wert, der automatisch ermittelt wird, vor. Die in dieser Arbeit für die qPCR verwendeten Primer wurden nach Möglichkeit in kodierenden Regionen und mit den in Tabelle 5 aufgelisteten Parametern mit Hilfe der Software Primer3 entworfen (s.o.). Entsprechend wurden alle anderen Kriterien in Kapitel 5.4.1 zu berücksichtigen. Alle Primer, die für die Validierung der CNVs genutzt wurden, sind in Appendix A2 aufgelistet.

Tabelle 5 Parameter in Primer3 zur Auswahl der qPCR Primer

	Minimum	Optimum	Maximum
PCR Produktgröße	120bp		150bp
Länge der Primer	18bp	20bp	24bp
Schmelztemperatur	59°C	60°C	61°C
GC-Kontent	20%	40%	60%

Alle Primerpaare wurden mit einer Kontroll-DNA anhand einer PCR und Gelelektrophorese überprüft. Die qPCR wurde in einer 384er Mikrotiterplatte mit einem minimalen Reaktionsansatz von 10µl durchgeführt. In Tabelle 6 ist der qPCR-Reaktionsansatz dargestellt. Die Tabelle 7 zeigt das qPCR-Programm mit SYBR Green I im LightCycler 480 II (Roche).

Tabelle 6 qPCR-Reaktionsansatz mit SYBR Green

	Menge
DNA (10ng/ μ l)	2 μ l
Primer F (10pmol)	0,4 μ l
Primer R (10pmol)	0,4 μ l
SYBR Green I Master (Roche)	5 μ l
Destilliertes Wasser	Ad 10 μ l

Tabelle 7 qPCR-Programm mit SYBR Green I (Roche)

Reaktionsphase	Temperatur	Zeit	Zyklen
initiale Denaturierung	95°C	10 min	1
Amplifikation	95°C	15 sec	40
	60°C	10 sec	
	72°C	15 sec	
Schmelzkurve	95°C	5 sec	1
	65°C	1 min	
	97°C	-	
Abkühlung	40°C	unendlich	1

Es ist besonders wichtig, spezifische Primer auszuwählen, um die Bildung von Primer-Dimeren zu vermeiden, da sich SYBR Green I unspezifisch in jede doppelsträngige DNA einlagert. Eine Schmelzkurve wurde analysiert, um die Spezifität der PCR-Reaktion zu überprüfen. Die Temperatur wird langsam kontinuierlich erhöht, dadurch wird die DNA bis zum spezifischen Schmelzpunkt aufgeschmolzen und dabei SYBR Green I freigesetzt. Dadurch wird eine Änderung der Fluoreszenz gemessen. Die Schmelzkurve und Schmelztemperatur sind abhängig von der Länge und Basenzusammensetzung des DNA-Abschnittes. Mittels der Schmelzkurve unterscheiden sich die unterschiedlichen PCR-Produkte bezüglich ihres Schmelzverhaltens. Die Peaks der Schmelzkurven stellen die Eigenschaft eines jeweiligen PCR-Produktes dar. Eine zusätzliche Schmelzkurvenspitze deutet darauf hin, dass unspezifische Nebenprodukte in der PCR-Reaktion entstehen.

Mittels der $\Delta\Delta C_t$ -Methode, die von Alexander Zink in unserem Labor etabliert wurde, wurden die Auswertungen der Daten durchgeführt. Grundsätzlich ist eine genaue Handhabung

essentiell bei der qPCR. Geringe Unterschiede der Konzentration von Patienten- und Kontrollpersonen-DNA können eine Verfälschung der qPCR-Ergebnisse verursachen. Außerdem kann man Pipettierfehler und Konzentrationsunterschiede nicht ganz vermeiden. Um die Ergebnisse zu kalibrieren, wurden daher endogene Kontrollen, die sogenannten Haushaltsgene (engl. housekeeping genes, HKGs), verwendet. Diese sollten stets in zwei Kopien im Genom vorhanden sein und ubiquitär exprimiert werden. Daher wurde der Ct-Wert der Zielsequenz gegen den Ct-Wert der internen Kontrollen in der $\Delta\Delta\text{Ct}$ -Methode normalisiert. ($\Delta\text{Ct} = \text{Ct}_Z - \text{Ct}_{\text{HKGs}}$). Als interne Kontrollen wurden in dieser Arbeit die drei Haushaltsgene *BNC1*, *CFTR* und *RnaseP* verwendet. Alle Messwerte wurden gegen diese drei internen Kontrollen normalisiert. Darüber hinaus wurde der ΔCt der Kontrollpersonen vom ΔCt der Patienten subtrahiert. ($\Delta\Delta\text{Ct} = [(\text{Ct}_Z - \text{Ct}_{\text{HKGs}})_{\text{Patient}} - (\text{Ct}_Z - \text{Ct}_{\text{HKGs}})_{\text{Kontrollperson}}]$). Nach der Normalisierung mit einer internen Kontrolle und Kontrollpersonen wurde die Anzahl an DNA-Zielabschnitten in der PCR ($2^{-\Delta\Delta\text{Ct}}$) ermittelt. Die relative Kopienzahl des DNA-Zielabschnittes im Genom der Patienten hinsichtlich der Kontrollpersonen ergab die Multiplikation $2 \times 2^{-\Delta\Delta\text{Ct}}$ und ist nicht absolut zu interpretieren, da die relative Kopienzahl mit dieser $\Delta\Delta\text{Ct}$ -Methode zwei betrug, obwohl die Patienten als auch Kontrollperson mehr oder weniger als zwei Kopien eines DNA-Abschnittes hatten. In dieser Arbeit wurde jeder Patient mit vier Kontrollpersonen verglichen, um daraus entstandene mögliche Fehler zu beseitigen.

5.4.5 Sanger-Sequenzierung von DNA-Fragmenten

Die Sanger-Sequenzierung, die sogenannte Kettenabbruch-DNA-Sequenzierung, ist nach wie vor dem Gold-Standard bezogen auf die verschiedenen DNA-Sequenzierungstechniken und wurde von Sanger et al. (1977) entwickelt. Mittels Sanger-Sequenzierung wird die Reihenfolge der Nukleotide in einem DNA-Molekül ermittelt (Sanger et al., 1977). Die wichtigste Reaktion ähnelt einer Standard-PCR-Reaktion mit dem Hauptunterschied, dass zusätzlich zu dNTPs fluoreszenzmarkierte Didesoxynukleotide (ddNTPs) hinzugefügt werden, welche für jede der vier Basen mit einem unterschiedlichen fluoreszierenden Farbstoff markiert sind. ddNTPs fehlt am C3-Atom der Ribose die Hydroxy-Gruppe, wodurch die Bindung eines weiteren dNTPs nicht mehr möglich ist und die Polymerisierungsreaktion abbricht. Da im Reaktionsansatz auch dNTPs enthalten sind, entstehen bei der Reaktion DNA-Stränge verschiedener Länge, deren

3'-terminale Enden mit dem fluoreszenzmarkierten ddNTP später mit einem Laser detektiert werden können. Bei einer Kettenabbruchreaktion wird jeweils nur ein Primer zugegeben, sodass die Reaktion nur in einer Richtung stattfindet. Die neu synthetisierten DNA-Fragmente wurden mittels des automatischen Kapillar-Sequenziergerätes 3130xl Genetic Analyzer (Life Technologies Corporation) sequenziert.

Das PCR-Produkt diente als Ausgangsmaterial und wurde zunächst nach Herstellervorgaben mit dem Agencourt AMPure XP-Kit (Beckman Coulter), der auf der SPRI-Technologie (Solid Phase Reversible Immobilization) basiert, aufgereinigt. In Abbildung 7 ist das Aufreinigungsverfahren mit dem AMPure Kit dargestellt. Nach der Aufreinigung des PCR-Produktes wurde die Sequenzierungs-PCR mit einem der PCR-Primer durchgeführt.

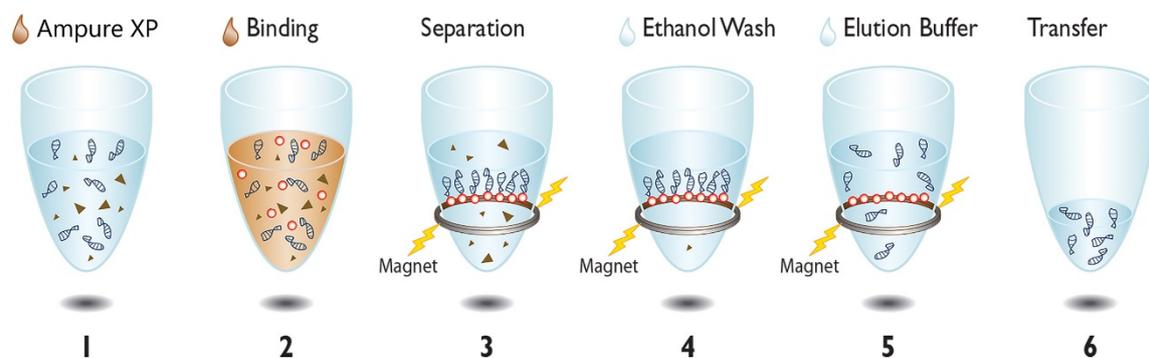


Abbildung 7 Aufreinigungsverfahren mittels Agencourt AMPure XP-Kit. 1. Zugabe von AMPure XP Reagenz zuden PCR-Produkten; 2. Bindung der PCR-Produkte an Magnet-Beads; 3. Trennen der PCR-Produkte von den Verunreinigungen mittels eines Magnetfeldes; 4. Waschen mit 70%-igem Ethanol; 5. Eluieren der PCR-Produkte mit TE⁻⁴-Puffer von den magnetischen Partikeln; 6. Überführung der PCR-Produkte in ein neues Tube (AMPure XP Protocol_000387v001).

Die aufgereinigten PCR-Produkte wurden als Ausgangsmaterial für die Sequenzierungs-PCR verwendet. Die in der Tabelle 8 aufgeführten Komponenten wurden für einen 20µl Sequenzierungsansatz eingesetzt. Anschließend wurde die Sequenzierungs-PCR im gleichen Thermozykler wie die PCR durchgeführt. Das Programm der Sequenzierungs-PCR ist in Tabelle 9 dargestellt.

Tabelle 8 Ansatz der Sequenzierungs-PCR

Komponenten	Menge
5xSequenzierungspuffer	3,75µl
BigDye 3.1 Mastermix	0,5µl
Primer F/R (3,2 pmol)	1µl
PCR-Produkt	1µl
Destilliertes Wasser	Ad 20µl

Tabelle 9 Standard-Programm der Sequenzierungs-PCR

Reaktionsphase	Temperatur	Zeit	Zyklen
initiale Denaturierung	95°C	1 min	1
Denaturierung	95°C	10 sec	25
Annealing	50°C	5 sec	
Elongation	60°C	4 min	
finale Elongation	72°C	10 min	1
Abkühlung	4°C	unendlich	1

Schließlich wurden die Sequenzierprodukte nach Herstellerangaben mit dem Agencourt CleanSEQ Kit (Beckman Coulter) aufgereinigt. Das Verfahren zur Aufreinigung der Sequenzierprodukte mittels eines CleanSEQ Kit ist schematisch in der Abbildung 8 dargestellt. In dieser Arbeit wurden 6µl aufgereinigtes Sequenzierprodukt mit 6µl HPLC-Wasser verdünnt und die Sequenz des Sequenzierproduktes wurde mittels eines 16-Kapillar-Sequenzierautomaten (3130xl Genetic Analyzer; Life Technologies Corporation) ermittelt, und anschließend mithilfe der Software SeqMan (DNASTAR) ausgewertet.

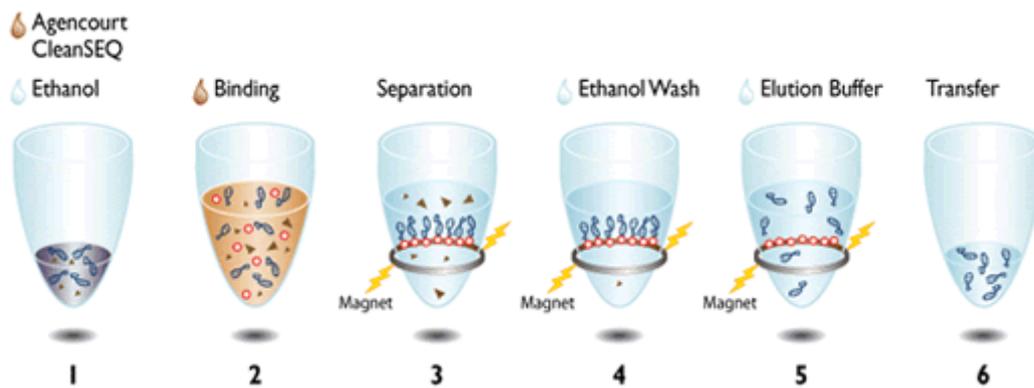


Abbildung 8 Verlauf der Aufreinigung des Sequenzierproduktes mittels eines CleanSEQ Kit. 1. Zugabe von CleanSEQ-Reagenz und 85%-igem Ethanol zu den Sequenzierprodukten; 2. Binden der Sequenzierprodukte an Magnet-Beads; 3. Trennen der Sequenzierprodukte von den Verunreinigungen mittels eines Magnetfeldes; 4. Waschen mit 85%-igem Ethanol; 5. Eluieren der Sequenzierprodukte mit HPLC-Wasser von magnetischen Partikeln; 6. Überführung der Sequenzierprodukte in ein neues Tube (CleanSEQ Protocol_000600v032).

5.4.6 Genotypisierung mittels Sequenom

Die SNP-Genotypisierung wurde über den iPLEX® Gold Assay auf der MassARRAY-Plattform von Sequenom durchgeführt. Mittels der MALDI-ToF MS Technologie (Matrix Assisted Laser Desorption/Ionization-Time of Flight Mass Spectrometry) und des iPLEX Gold Kit wurde die Analyse einer Vielzahl von ausgewählten Markern in allen Proben ermöglicht. Die Methode erlaubt es, bis zu 40 SNPs pro Multiplex-Reaktion in 384 Patientenproben parallel zu analysieren. So konnte die Replikationsstudie zu Kandidaten-Loci durchgeführt werden. Grundsätzlich liegt das Prinzip dieser Technologie in der Differenzierung von Allelen aufgrund der Massenunterschiede der vier Basen (Haff and Smirnov, 1997).

Zunächst wird das Plexdesign (Primerdesign) zur Amplifikation und Einzelbasen-Extension durch die von Agena Bioscience bereitgestellte Assay Design Suite Software durchgeführt. Nach Eingabe von SNP-ID oder Informationen zur Markersequenz wird eine Zusammenstellung der Primersequenzen für jeden einzelnen ausgewählten Marker von einem Algorithmus dieser oben genannten Software ermittelt. Für das Primerdesign müssen verschiedene Parameter, einschließlich der Primerdimerisierung, der Schmelztemperatur, sowie der in der Primerbindungsstelle liegenden bekannten Varianten, berücksichtigt werden. Die Anzahl der Primerpaare entspricht der Anzahl der zu analysierenden Marker.

Im ersten Schritt wird durch eine Multiplex-PCR mit den für jeden SNP spezifischen PCR-Primern ein ca. 100bp großes PCR-Produkt generiert. Die Reaktion wurde in 384er Mikrotiterplatten angesetzt. Für ein Sequenom-Experiment wurden in dieser Arbeit 16ng DNA als Ausgangsmaterial verwendet. Diese DNA wurde zuerst in 384er Mikrotiterplatten mittels Pipettier-Roboter vorgelegt und dann im Trockenschrank bei Raumtemperatur über Nacht eingetrocknet. Danach wurden 5µl des PCR-Reaktionsansatzes in die Platten hinzu pipettiert und mit einem spezifischen Programm wurde die PCR in einem Themozykler durchgeführt. Der Reaktionsansatz und das PCR-Programm sind in den Tabellen 10 und 11 dargestellt.

Tabelle 10 Reaktionsansatz der Multiplex-PCR

Komponente	Menge
10XPuffer (inkl. 15mM MgCl ₂)	0,625µl
dNTPs (10mM)	0,25µl
MgCl ₂ (25mM)	0,325µl
Primer-Mix (Plex, 0,5µM)	1µl
HotStar Taq DNA-Polymerase (5U/µl)	0,1µl
DNA	16ng (eingetrocknet)
Destilliertes Wasser	Ad 5µl

Tabelle 11 Temperaturprogramm einer Multiplex-PCR

Reaktionsphase	Temperatur	Zeit	Zyklen
initiale Denaturierung	95°C	15 min	1
Denaturierung	95°C	20 sec	45
Annealing	56°C	30 sec	
Elongation	72°C	1 min	
finale Elongation	72°C	3 min	1
Abkühlung	12°C	unendlich	1

Nach der Multiplex-PCR wurden die PCR-Produkte anschließend mit Schrimp Alkaline Phosphatase (SAP) verdaut, welche die überschüssigen dNTPs entfernte, die sonst bei der Einzelbasen-Verlängerungsreaktion (engl. single-base extension, SBE) stören könnten. Der

Reaktionsansatz und das Temperaturprogramm des SAP-Verdau sind in den Tabellen 12 und 13 dargestellt.

Tabelle 12 Reaktionsansatz des SAP-Verdau

Komponente	Menge
Multiplex-PCR Produkt	5µl
10XSAP-Puffer	0,17µl
SAP (1U/µl)	0,3µl
Destilliertes Wasser	Ad 5µl

Tabelle 13 Temperaturprogramm des SAP-Verdau

Reaktionsphase	Temperatur	Zeit
Verdau	37°C	40 min
Inaktivierung	85°C	5 min
Abkühlung	12°C	unendlich

Nach dem SAP-Verdau wurde die Einzelbasen-Verlängerungsreaktion (engl. single-base extension, SBE) durchgeführt, indem der UEP-Mix (engl. unextended primer, UEP) und Didesoxynukleotide (ddNTPs) auf die 384-er Platte gegeben wurden. Anhand der Assay Design Suite Software wurden ebenfalls die UEPs ermittelt, die alle eine unterschiedliche Masse haben. Die UEPs hybridisierten mit ihrem Hydroxyl-Ende direkt an die erste Position neben dem SNP, wodurch es zur Verlängerung mit einem ddNTP und zu den Genotyp-spezifischen Massendifferenzen kommt. Die ddNTPs dienten ähnlich der Sequenzier-Reaktion in der Sanger-Sequenzierung als Terminatoren, dadurch, dass die Hydroxylgruppe nicht mehr an eine Carboxylgruppe bindet. So wurde die Reaktion nach der Verlängerung von einer Base nach dem SNP abgebrochen. Für die SBE-Reaktion wurden die UEPs aufgrund der unterschiedlichen Massengruppen in vier unterschiedlichen Konzentrationsgruppen (7µM, 9,3µM, 11,6µM, 14µM) eingeteilt. Die Sequenzen der in dieser Arbeit verwendeten Vorwärts-, Rückwärts- und UEP-Primer, sowie die im Plex vorhandenen Komponenten, finden sich in

Appendix A2. Der Reaktionsansatz für die SBE und das verwendete Temperaturprogramm sind in Tabelle 14 und Tabelle 15 dargestellt.

Tabelle 14 Reaktionsansatz der SBE-Reaktion

Komponente	Menge
Multiplex-PCR Produkt (nach SAP-Verdau)	7µl
UEP-Mix	0,94µl
10XiPLEX Gold Puffer	0,2µl
iPLEX Termination Mix	0,2µl
iPLEX Gold Enzym	0,041µl
Destilliertes Wasser	Ad 9µl

Tabelle 15 Temperaturprogramm der SBE-Reaktion

Reaktionsphase	Temperatur	Zeit	Zyklen	
initiale Denaturierung	94°C	30 sec	1	
Denaturierung	94°C	5 sec	1	45
Annealing	56°C	5 sec	5	
Elongation	72°C	5 sec		
finale Elongation	72°C	3 min	1	
Abkühlung	12°C	unendlich	1	

Um störende Kationen aus dem Reaktionsansatz für die massenspektrometrische Analyse zu beseitigen, wurde ein Ionenaustauscher-Harz (clear resin) benutzt (Oeth et al., 2005). 6mg des Harzes wurden in jede Vertiefung einer speziell dafür vorgesehenen 384er Platte ausgestrichen und für mindestens eine Stunde getrocknet. Zuerst wurden je 16µl Millipore-Wasser zur Reaktion hinzupipettiert. Danach wurde die Reaktion durch Zugabe von trockenem Harz über einen Ionenaustausch entsalzt. Hiernach wurde der Reaktionsansatz für sieben Minuten langsam rotiert und anschließend bei 4000 rpm für sieben Minuten zentrifugiert. Danach wurden 15nl der jeweiligen Multiplex-Reaktion aus der Platte mittels eines Nanodispensers unter Verwendung der SpectroPoint Software auf einen SpectroCHIP

überführt. Auf dem Chip sind genau 384 Matrixspots mit Hydroxypicolinsäure besetzt. Zusätzlich wird ein Kalibrant gespottet, der drei Oligonukleotide bekannter Größe enthält, um das Gerät initial zu kalibrieren. Anschließend wurde die MALDI-ToF-MS-Analyse im Sequenom-Analysegerät durchgeführt.

Ein kurzer Impuls durch einen UV-Laser unter Vakuum ionisiert dabei jede Probe und führt dazu, dass sich ionisierte Proben aus der Matrix lösen. Die ionisierten Proben werden durch ein elektrisches Feld beschleunigt und wandern im Flugrohr Richtung Kathode. Ihre Flugzeiten (engl. time of flight, TOF) sind abhängig von ihren Massen und werden dabei durch den Ionendetektor erfasst. Die Analyse der Probe mit einer Masse zwischen 4.500 und 9.000 Dalton wurde über den MassARRAY Compact Analyzer durchgeführt, welcher für jede Probe ein Massenspektrum erstellt.

5.4.7 Genotypisierung mittels Illumina

Außer Sequenom wurde die Infinium BeadArray Plattform der Firma Illumina für die Genotypisierung verwendet. Mittels Illumina Arrays erlauben genomweite Assoziationsstudien und die Detektion von Kopienzahlveränderungen. Hierbei ermittelt man durch eine allelspezifische Einzelbasenverlängerung den Genotyp des einzelnen SNP (Steemers et al., 2006). Um zu genotypisieren, wurden BeadChips verwendet. Auf den Chips sind große Siliziumkugeln, die sogenannten Beads, mit einer Größe von 3µm. Der BeadArray basiert auf einer direkten Hybridisierung von genomischer DNA der Patienten an 50bp lange, markerspezifische Sonden (Steemers et al., 2006). Auf dem Chip befindet sich jeder Bead an einer bestimmten Stelle und an jeden Bead koppelt nur die Sonde für einen spezifischen Marker (SNP). Ein SNP-spezifischer Bead ist im Mittel 20mal auf einem Chip vorhanden (Gunderson et al., 2004). Jeder Bead ist zusätzlich zur jeweiligen SNP-spezifischen Sonde mit einer einmaligen Sequenz zur Identifikation gekennzeichnet. Die Auswahl der SNPs wurde unter Berücksichtigung des HapMap-Projektes und dem 1.000-Genome-Projekt von Illumina vorgenommen.

Die Genotypisierung mittels Illumina Arrays wurde nach Herstellervorgaben durchgeführt und bestand aus vier Schritten: (i) Amplifikation des ganzen Genoms, (ii) Array-basierte Hybridisierungserfassung, (iii) enzymatische allelspezifische Primerverlängerung (engl.

enzymatic allele-specific primer extension, ASPE), (iv) Detektion des amplifizierten Signals. Insgesamt dauert dieses Verfahren drei Tage.

Zunächst wurde eine Verdünnung von 50ng/μl (insgesamt 500ng) der genomischen DNA jedes Patienten als Vorbereitung hergestellt. Am ersten Tag wurde die Amplifikation der DNA durchgeführt, um qualitativ hochwertige Genotypisierungsdaten zu bekommen. Am zweiten Tag wurde zunächst die amplifizierte DNA enzymatisch fragmentiert, anschließend gefällt und neu gelöst. Schließlich wurde die behandelte DNA an die Sonden der Beads spezifisch hybridisiert. Die Sequenzen der Sonden unterscheiden sich jeweils nur in der letzten Base vor dem zu untersuchenden Marker. Am dritten Tag erfolgten die letzten zwei Schritte. Die BeadArrays wurden zuerst gewaschen und eine Einzelbasenverlängerung mit einer Färbung wurde durchgeführt. Anschließend wurden die fertig prozessierten Chips mittels des iScan Scanners von Illumina ausgelesen.

Ein Antikörper Sandwich-Assay wurde durchgeführt, damit die zu messende Fluoreszenz der SNPs verstärkt wurde. Die ddNTPs waren mit zwei verschiedenen Chemikalien markiert: ddCTPs und ddGTPs mit Biotin, ddATPs und ddUTPs mit DNP (2,4-dinitrophenol). Die Hapten-markierten ddNTPs wurden effizient durch Polymerase eingebaut und erlaubten den Nachweis mit einem zweifarbigen, orthogonalen, mehrschichtigen immunhistochemischen Sandwich-Assay (Steemers et al., 2006). Biotin und DNP wurden mit einer Kombination von Alexa555-markiertem Streptavidin (SA) und Alexa647-markiertem Primärantikörper aus Kaninchen gegen DNP angefärbt. Des Weiteren erfolgte die Gegenfärbung mit biotinyliertem Anti-SA und DNP-markiertem Ziegen-anti-Kaninchen-Sekundärantikörper. Um eine optimale Sensitivität zu erhalten, wurde eine weitere Runde der Sandwichfärbung für Biotin und DNP durchgeführt. Zum Schluss wurden die BeadArrays mittels eines iScan Scanner von Illumina ausgelesen und mittels der GenomeStudio Software wurden die Genotypisierungsdaten ausgewertet.

In der vorliegenden Arbeit wurden die HumanOmniExpressExome-8 v1.2 und v1.3 Arrays verwendet. Der HumanOmniExpressExome-8 v1.3-BeadArray enthält zusätzlich 4.854 Marker im Vergleich zum HumanOmniExpressExome-8 v1.2-BeadArray, da die SNP-Arrays kontinuierlich weiterentwickelt werden. Die neuere Generation von BeadArrays enthält jeweils eine größere Anzahl Marker.

5.4.8 Whole-Exome Sequencing (WES)

Das Whole-Exome Sequencing (WES) besteht aus Erfassung, Sequenzierung und Analyse aller Exons aller Protein-kodierenden Gene im menschlichen Genom. Anstatt des gesamten Genoms konzentriert sich WES nur auf die etwa 30 Millionen Basenpaare, die in funktionelle Proteine umgesetzt werden, in denen Mutationen aller Wahrscheinlichkeit nach eine schwere direkte phänotypische Konsequenz haben (Bertier et al., 2016). In der vorliegenden Arbeit wurden WES von 30 Trios mit EA/TEF mit Hilfe des CCG Köln (Cologne Center for Genomics) durchgeführt.

Exonische und benachbarte intronische Sequenzen wurden aus genomischer DNA mit dem NimbleGen SeqCap EZ Human Exome Library v2.0 angereichert. WES wurde unter Verwendung eines 2x100bp-Paired-End-Leseprotokolls (Protokoll v1.2) aufgrund der Empfehlungen des Herstellers mittels eines Illumina HiSeq2000-Sequenzers in Köln am CCG durchgeführt. Die primären Rohdaten wurden nach Signalreinheit durch die Illumina Realtime Analysis (RTA) Software v1.8 gefiltert. Anschließend wurden die Reads auf das menschliche hg19-Referenzgenom mittels des Algorithmus bwa-aln abgebildet (Li and Durbin, 2009; Zhang et al., 2016). Die duplizierten Reads wurden mit Hilfe der Software GATK v1.6 für die lokale Neuausrichtung auf kurze Insertionen und Deletionen markiert, die Qualitätswerte der Basen neukalibriert, sowie die SNVs (engl. single-nucleotide variants) und Indels (engl. insertions and deletions) detektiert (Kawalia et al., 2015; McKenna et al., 2010). Die Skripte, die im CCG Köln entwickelt wurden (unveröffentlicht), wurden verwendet, um die Allelfrequenzen, die von der ESP6500-Datenbank [Exome Variant Server, NHLBI GO Exome Sequencing Project (ESP), Seattle, WA (<http://evs.gs.washington.edu/EVS/>)] berichtet wurden, zu erfassen und die Veränderungen in der Proteinstruktur zu detektieren. Mutationen in den Akzeptor- und Donor-Spleisstellen wurden mit einem Maximum-Entropy-Modell analysiert (Yeo and Burge, 2004). Mittels der Software DeNovoGear v.0.5.1 wurden die potenziellen *de novo* Mutationen in den Trios untersucht (Ramu et al., 2013; Zhang et al., 2016).

5.4.9 RNA-Sequenzierung (RNA-Seq)

In dieser Arbeit wurde die RNA-Sequenzierung für die Transkriptom-Analyse mit Hilfe des QuantSeq 3' mRNA-Seq Library Prep Kit (Lexogen) nach den Herstellerangaben durchgeführt. Hierbei wurde für jedes Transkript nur ein einzelsträngiges spezifisches Fragment zur Sequenzierung am 3'-Ende der PolyA-RNA generiert.

Als Ausgangsmaterial wurden 200ng gesamt-RNA von Mäusen verwendet und eine Amplifikation mit 13 PCR-Zyklen durchgeführt. Die Qualitätskontrolle erfolgte mittels einer TapeStation2200 (Agilent), um zu überprüfen, ob alle Proben die minimale RNA-Konzentration von 2nmol/l überschritten hatten. Die Sequenzierung wurde im HiSeq2500 (Illumina) auf dem Rapid-Run-Modus mit V2 Chemikalien (Illumina) und zwei technischen Replikaten der Probe durchgeführt. Nach dem Demultiplexen mittels bcl2fastq (Illumina) wurde die webbasierte Software FastQC v0.11.4 (<http://www.bioinformatics.babraham.ac.uk/projects/>) verwendet, um die Qualität der FASTQ-Daten zu kontrollieren. Das Alignment der Reads erfolgte mithilfe des Algorithmus Subread v1.5.1 (<http://subread.sourceforge.net/>) gegen die primäre Assemblierung des murinen Referenzgenoms GRCm38 (Liao et al., 2013). Die Zählung der Reads wurde anschließend mittels featureCounts v1.5.1 und der Ensembl-Gen-Annotation (Release 88) durchgeführt (Liao et al., 2014).

5.5 Bioinformatische Methoden

5.5.1 GenomeStudio (GS)

Um die Daten, die mittels Illumina SNP-Arrays generiert sind, zu visualisieren und analysieren, steht GenomeStudio 2.0.2 (Genotyping module 2.0.2, Illumina) zur Verfügung. Mithilfe des Genotypisierungsmoduls im GenomeStudio werden Analysen hinsichtlich des Verlustes der Heterozygotie (engl. loss of heterocycosity, LOH) und von Kopienzahlveränderungen (CNVs) möglich.

Für jeden SNP-Marker ergibt das zweifarbige Auslesen des Infinium HD Assays die Intensitätswerte, die in jedem der beiden Farbkanäle (zwei Allele, A-Allel und B-Allel) gemessen werden. Das A-Allel entspricht den rot markierten Nukleotiden A oder T, während das B-Allel die Nukleotide C oder G grün erfasst. Die Polartransformation dieser Daten liefert

normalisierte Intensitätswerte[®] und allele Intensitätsverhältnisse (Θ) (Guo et al., 2014). Diese Parameter können im GenomeStudio als Genoplot visualisiert werden. Darüber hinaus werden solche Werte verwendet, um zwei Metriken für jeden SNP-Marker in einer Probe zu berechnen, damit die Genotypen der SNPs ermittelt und die Kopienzahlen abgeschätzt werden können. Mittels GenomeStudio werden Plots aller SNPs für die B-Allel Frequenz (BAF) und das Log R Ratio (LRR) generiert, und so die Kopienzahlveränderungen ermittelt.

B-Allel Frequenz (BAF)

Die BAF zeigt die Genotypinformation eines untersuchten Markers einer Person und ermittelt den Thetawert für jeden Marker einer Probe, welcher für die Position dieser Probe innerhalb des Clusters korrigiert ist. Die Positionen innerhalb des Clusters wurden von einer großen Menge gesunder Individuen generiert. Mithilfe der BAF ist es einfacher, verglichen mit dem Thetawert, die Genotypdaten aller SNPs innerhalb einer chromosomalen Region zu visualisieren, da die BAF weniger Locusvariationen als der Thetawert für eine untersuchte Probe darstellt. Die Transformation des Thetawertes zu den Allelfrequenzen (Tabelle 16) ermöglicht verbesserte Messungen und eine bessere Visualisierung der Kopienzahlveränderungen.

Tabelle 16 Transformierung des Thetawertes zur BAF

BAF	Bedingungen
0	Wenn Thetawert < tAA
$0,5 * (\text{Thetawert} - tAA) / (tAB - tAA)$	Wenn Thetawert < tAB
$0,5 + 0,5 * (\text{Thetawert} - tAB) / (tBB - tAB)$	Wenn Thetawert < tBB
1	Wenn Thetawert \geq tBB

tAA: durchschnittlicher Thetawert aller Proben mit dem Genotyp AA im Cluster nach der Normalisierung;
tAB: durchschnittlicher Thetawert aller Proben mit dem Genotyp AB im Cluster nach der Normalisierung;
tBB: durchschnittlicher Thetawert aller Proben mit dem Genotyp BB im Cluster nach der Normalisierung.

Die BAF variiert zwischen 0 und 1. Bei Homozygotie zeigt die BAF entweder 0 oder 1. Bei Autosomen kann die BAF 0, 0,5 oder 1 betragen. Eine intermediäre BAF tritt bei einem Allel-Ungleichgewicht auf. Mittels BAF wird die relative Intensitätsqualität eines Allels im Vergleich zum anderen Allel indiziert (Attiyeh et al., 2009).

Log-R-Ratio (LRR)

Das LRR ($\log_2(R_{\text{Gemessen}}/R_{\text{Referenz}})$) basiert auf den normalisierten Intensitätsdaten und bezeichnet das Verhältnis des \log_2 -Wertes des normalisierten R-Wertes eines SNP zu dem normalisierten R-Wert einer Referenzprobe von Illumina (Peiffer et al., 2006). Ein LRR von 0 indiziert, dass die gemessene Intensität einer Probe gleich dem Wert der Referenzprobe ist. Ein LRR kleiner (größer) als 0 impliziert eine niedrigere (höhere) Intensität der gemessenen Probe verglichen zur Referenzprobe von Illumina.

CNV Detektion in GenomeStudio

Mittels der obengenannten Parameter BAF und LRR ermittelt man die Kopienzahl eines DNA-Abschnittes der zu untersuchenden Probe. Auf den Autosomen liegen die Gene in zwei Kopien im Genom vor. Beträgt die Kopienzahl eines Genabschnittes einer Probe zwei im Vergleich zur Referenzprobe, liegen die BAF bei 0, 0,5 oder 1 und es ergibt sich ein LRR-Wert von null. Die Abbildung 9 zeigt exemplarisch eine Darstellung von Chromosom 3 eines Patienten ohne Kopienzahlveränderungen im GenomeStudio. In dieser Studie erfolgte keine Analyse der Geschlechts-Chromosomen.

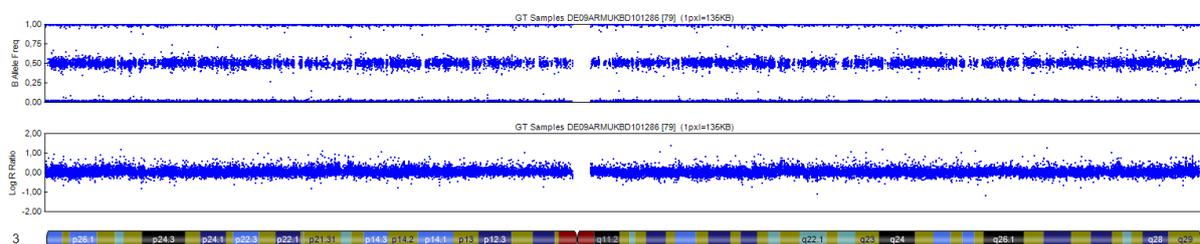


Abbildung 9 Beispiel für die Darstellung von Chromosom 3 eines Patienten ohne Kopienzahlveränderungen im GenomeStudio (oberer Teil: BAF, unterer Teil: LRR). Jeder blaue Punkt stellt einen Marker da. Das hier dargestellte LRR liegt bei einem Wert von null und deutet darauf hin, dass die gemessene Intensität eines Markers einer Probe identisch zur gemessenen Referenzprobe ist. Die BAF zeigt hierbei drei verschiedene mögliche Genotypen und weist um 0.0 oder 1.0 den Genotyp AA oder BB und um 0.5 den Genotyp AB auf.

Eine reduzierte Kopienzahl eines DNA-Abschnittes und reduzierte Intensität der analysierten Marker weisen auf eine Deletion in einer Probe hin, wobei das LRR kleiner als 0 ist. Die BAF bei einer heterozygoten Deletion beträgt entweder 0 oder 1 (Abb.10). Bei einer homozygoten Deletion zeigt sich eine deutlich reduzierte Intensität aller Marker und ein noch kleineres LRR. Theoretisch dürfte keine BAF für diesen DNA-Abschnitt vorliegen, allerdings weist die BAF zufällig 0 oder 1 bei der homozygoten Deletion auf, weil die markerspezifischen Beads teilweise ohne hybridisierte DNA eingefärbt werden und es so zu einem Artefakt kommt.

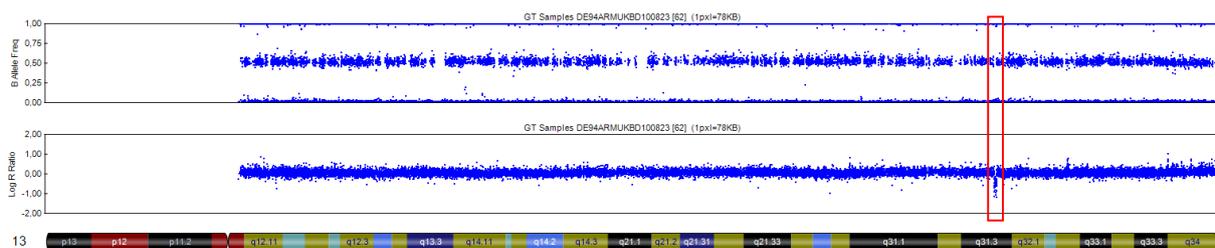


Abbildung 10 Beispiel für die Darstellung von Chromosom 13 eines Patienten mit einer heterozygoten Deletion im GenomeStudio (oberer Teil: BAF, unterer Teil: LRR). Jeder blaue Punkt repräsentiert einen Marker. Die kleine heterozygote Deletion ist rot markiert.

Bei einer Duplikation erhöht sich die Anzahl des DNA-Abschnittes einer Probe verglichen mit einer Referenz und das LRR (größer als 0) steigt durch eine erhöhte Intensität der in der Region liegenden Marker. Abbildung 11 zeigt ein Beispiel für eine Duplikation eines Chromosomenabschnittes von einem Patienten, wobei die BAF resultierend aus einer theoretischen Verteilung der zu erwartenden Genotypen um 0 oder 1 für die Genotypen AAA und BBB, um 0,33 und 0,67 für die Genotypen AAB und ABB aufweist. Bei einer Multiplikation ist die Kopienzahl deutlich größer als bei einer Duplikation und weist mehrere Kombinationen der Genotypen mit entsprechenden BAFs auf.

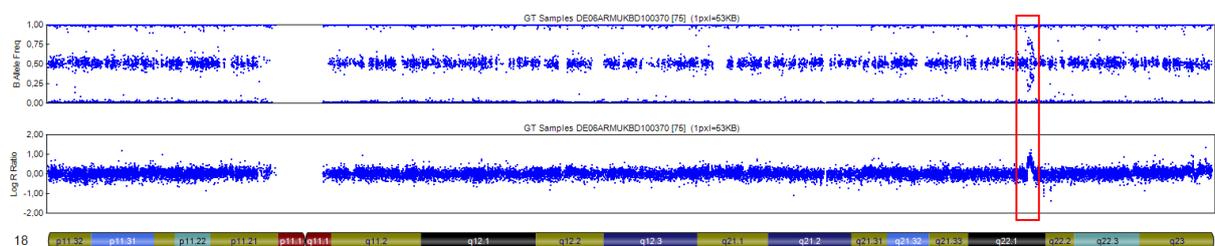


Abbildung 11 Beispiel für die Darstellung von Chromosom 18 eines Patienten mit einer Duplikation (rotmarkiert) im GenomeStudio (oberer Teil: BAF, unterer Teil: LRR). Jeder blaue Punkt repräsentiert einen analysierten Marker.

5.5.2 Detektion der CNVs aus genomweiten SNP-Arrays mittels QuantiSNP

Es ist aufwendig, alle CNVs der untersuchten Patienten im GS visuell zu beurteilen. Außerdem lassen sich kleine CNVs als auch ihre exakte Größe nicht vollständig beurteilen. Daher war ein Computeralgorithmus für die systematische Identifizierung dieser CNVs erforderlich.

Bisher wurden verschiedene „Software Tools“ entwickelt, um die in mittels SNP-Arrays erhobenen CNVs zu detektieren. Basierend auf den zugrundeliegenden mathematischen Methoden lassen sich diese Programme in zwei Gruppen unterteilen, die entweder auf die Anwendung eines Hidden Markov Modells (HMM) oder der Nutzung eines Segmentierungsalgorithmus beruhen. HMM-basierten Methoden zielen darauf ab, verdeckte Kopienzustände entlang einer Markov-Kette vorherzusagen, während Segmentierungsalgorithmen die Chromosomen in Segmente aufteilen und versuchen jedem Segment einen Kopienzustand sinnvoll zuzuordnen (Nutsua et al., 2015).

In der vorliegenden Arbeit wurde der Algorithmus QuantiSNP verwendet, um die Kopienzahlveränderungen mittels Illumina SNP-Arrays zu identifizieren. QuantiSNP basiert auf dem „objektive Bayes Hidden Markov Modell“ (OB-HMM) und ist speziell für die Detektion der Regionen mit Kopienzahlveränderungen mittels BeadArrayTM Genotypisierungsdaten im Hinblick auf das LRR und die BAF jedes SNP ausgerichtet (Colella et al., 2007). Mittels QuantiSNP wird ein LogBayes-Faktor (logB) jeder Region mit der Kopienzahlveränderung zugeordnet. Der logB ermittelt eine Wahrscheinlichkeit für die Effektstärke aus den Daten für das Vorhandensein einer Kopienzahlvariante in einer Region verglichen zur Nullhypothese bzw. dass keine Variante in der Region vorliegt. Je größer der logB ist, desto stärker ist der Beweis für die Existenz einer Kopienzahlvariante. Die QuantiSNP-Version 2.0 (<https://sites.google.com/site/quantisnp/>) schlägt als Schwellenwert einen logB Wert von mindestens 30 vor, um eine möglichst niedrige Falsch-Positiv-Rate (<1%) zu erhalten. Konsekutiv erhöhen Schwellenwerte zwischen 10 und 30 die Wahrscheinlichkeit kleinere CNVs zu detektieren, allerdings unter Inkaufnahme einer steigenden Falsch-Positiv-Rate (bis zu 10%). Kopienzahlveränderungen mit einem logB kleiner als 10 sind fraglich und werden vom Programmhersteller nicht empfohlen. In dieser Arbeit wurde, basierend auf den Erfahrungen der eigenen AG, ein Schwellenwert von logB 30 festgesetzt, um so die Falsch-Positiv-Rate an detektierten CNVs möglichst gering zu halten.

5.5.3 Cartagenia Bench Lab CNV

Die Software Cartagenia Bench Lab CNV ist eine webbasierte Datenverwaltungsplattform, um die beobachteten strukturellen genomischen Varianten und phänotypische Daten zu speichern, zu visualisieren und zu interpretieren. Auf dem Server der Humangenetik in Bonn steht diese Software zur Verfügung. Die Software Cartagenia Bench Lab CNV ist mit verschiedenen Datenbanken vernetzt, einschließlich der Ensembl-Datenbank, der DGV-Datenbank usw., damit eine Verarbeitung der erhobenen QuantiSNP-Daten halbautomatisch durchgeführt werden kann. Nach dem Hochladen der QuantiSNP-Daten erlaubt das Programm die detektierten CNVs mit eigenen, dem biologischen Modell angepassten Filterkriterien weiter zu bearbeiten.

5.5.4 Filterkriterien der CNV-Analyse

Um bei der Analyse der CNVs falsch-positive Befunde zu vermindern und polymorphe CNVs zu entdecken, ist die Auswahl der Filterkriterien zwingend notwendig. In diesem Kapitel werden die Filterkriterien für die in dieser Arbeit durchgeführten CNV-Analysen beschrieben und in der Abbildung 12 dargestellt.



Abbildung 12 CNV-Filter Schritte von der Genotypisierung bis zur Verifikation.

Zunächst wurden die Genotypisierungsdaten mittels HumanOmniExpressExome-8 v1.2 und v1.3 von der Firma Illumina erhoben. Im Folgenden wurde eine Qualitätskontrolle im GS durchgeführt. Eine Call-Rate steht hier für die Qualität der Genotypisierungsdaten und wurde im GS für jede Probe ermittelt. Die Proben, die eine Call-Rate kleiner als 95% aufweisen, wurden von der Analyse der CNVs ausgeschlossen, da weniger als 95% der Marker erfolgreich für diese Proben genotypisiert waren. Nach der Qualitätskontrolle im GS wurde ein Endbericht für QuantiSNP generiert.

Bei der Filterung in QuantiSNP wurden der Wert logB und Anzahl der Marker berücksichtigt. Wie oben beschrieben, wurde der Wert logB um 30 festgesetzt, um die Falsch-Positiv-Rate zu reduzieren. Die Anzahl der CNV-beinhaltenen Marker wurde in dieser Arbeit mit 5

angenommen, d.h. um einen CNV beurteilen zu können, sollten mindestens fünf aufeinanderfolgende Marker sicher typisiert sein.

Die von QuantiSNP erhobenen CNV-Daten wurden in Cartagenia Bench Lab CNV für weitere Filterschritte hochgeladen. Hierbei wurden zunächst nochmal der Wert $\log B$ um 30 und die Anzahl der Marker auf mindestens fünf eingestellt. Des Weiteren wurden Geninhalt und Größe der CNVs beachtet, und schließlich wurden die erhobenen CNVs gegen HNR-Kontrollen gefiltert. Dabei kann es sein, dass dieselbe Anzahl an Markern physikalisch unterschiedlich große genomische Bereiche beschreiben. Dies ist dadurch zu erklären, dass manche Regionen eine große Anzahl an informativen Markern in enger physikalischer Nähe zueinander beinhalten und wiederum andere Regionen nur wenige informative Marker in enger physikalischer Nähe zueinander beinhalten. Ziel der CNV-Analyse war es, CNVs zu detektieren, die kodierenden Abschnitte, also Gene, beinhalten. Daher wurden die CNVs, in denen keine Gene vorliegen, ausgefiltert. Darüber hinaus wurden die zu untersuchenden CNVs mit den vorhandenen CNVs in Core DGV Raw Feb-13 verglichen (<http://dgv.tcag.ca/dgv/app/home>). Folgende CNVs wurden ausgefiltert: (i) CNVs, die komplett mit einem CNV aus DGV überlappten, (ii) CNVs die partiell mindestens dreimal mit in DGV aufgeführten CNVs überlappten, (iii) CNVs, die in Regionen kartierten, welche bereits als benigne bei Cartagenia Bench Lab CNV klassifiziert wurden. Nach dem Vergleich mit den CNVs in DGV wurden alle verbleibenden CNVs gegen die verfügbare Kontrollkohorte (HNR-Kontrolle) gefiltert. Nur die CNVs, die keine physikalische Überlappung mit den CNVs von HNR-Kontrollen aufwiesen, wurden für weitere Filterschritte und Analysen aufgenommen.

Alle verbleibenden CNVs wurden mittels UCSC Genombrowser und GS visuell inspiziert. Die CNVs, die nicht in kodierenden Regionen lagen und nur mit vereinzelt Markern ($n < 5$) abgedeckt waren, wurden in diesem Schritt ausgegrenzt.

Alle verbliebenen putativen CNVs wurden abschließend mittels qPCR validiert. Die verifizierten CNVs wurden hinsichtlich der Funktionen und Expressionen der Gene, die in den Regionen der CNVs vorliegen, sowie der bekannten krankheitserzeugenden CNVs in vergangenen CNV-Studien als pathogen oder benigne beurteilt.

5.5.5 Analyse der Exom-Daten mittels VARBANK

Für die Analyse der Exom/Genom-Daten wurde von Herrn Holger Thiele im CCG Köln die Software Varbank entwickelt, die durch einen Webbrowser erreichbar ist und die grundsätzlichen Annotationen sowie die Filterfunktionen zur Verfügung stellt. Daher wurden alle in der Arbeit analysierten Exom-Daten nach Absprache mit Herrn Thiele mittels der Varbank gefiltert und analysiert.

Die Varbank (unveröffentlicht, <https://varbank.ccg.uni-koeln.de>) wurde verwendet, um die Exom-Daten zu filtern. Dabei können durch die Varbank seltene ($MAF < 0,005$) Varianten oder *de novo* Varianten (die A-posteriori-Wahrscheinlichkeit einer *de novo* Mutation, engl. posterior probability of a *de novo* mutation, $PP_DNM > 0,5$) sowie die Varianten, welche die Proteinstruktur oder die Spleißstelle beeinflussen gefiltert werden. Außerdem wurden die Varianten gegen eine interne Datenbank gefiltert, die alle Varianten von 511 Exomen von Patienten mit Epilepsie enthält, um Pipeline-bezogene Artefakte auszuschließen. Die nach den Filterschritten verbleibenden Varianten wurden visuell in der Varbank inspiziert. Schließlich wurden alle Varianten, die alle Filterkriterien erfüllten, mittels Sanger-Sequenzierung validiert.

5.5.6 dbNSFP zur Interpretation der detektierten Varianten

dbNSFP ist eine Datenbank, die für die funktionale Vorhersage und die Annotation aller potentiellen nicht-synonymen Einzelnukleotidvarianten (engl. non-synonymous single-nucleotide variants, nsSNVs) im menschlichen Genom entwickelt wurde.

Die aktuelle Version von dbNSFP ist auf der GENCODE-Version 22 / Ensembl-Version 79 abgelegt und umfasst insgesamt 83.422.341 nsSNVs und ssSNVs (engl. splicing-site SNVs). Die Version beinhaltet Prädiktionswerte aus 20 verschiedenen Prädiktionsalgorithmen einschließlich SIFT, Polyphen2-HDIV, Polyphen2-HVAR, LRT, MutationTaster2, MutationAssessor, FATHMM, MetaSVM, MetaLR, CADD, VEST3, PROVEAN, FATHMM-MKL coding, fitCons, DANN, GenoCanyon, Eigen coding, Eigen-PC, M-CAP, REVEL und MutPred. Des Weiteren beinhaltet die aktuelle Version von dbNSFP 6 Konservierungswerte (PhyloP x 2, phastCons x 2, GERP++ und SiPhy) und weitere relevante Angaben einschließlich Allelfrequenzen, die in den Daten aus 1000 Genomes Projekt Phase 3, UK10K-Kohorten, ExAC Konsortium und NHLBI Exome Sequenzierungsprojekt ESP6500 beobachtet wurden. Zudem

beinhaltet die aktuelle Version von von dbNSFP verschiedene Gen-IDs aus unterschiedlichen Datenbanken, die Funktionsbeschreibung und Expressionsdaten von Genen, sowie Geninteraktionsinformationen usw. (Liu et al., 2011, 2016).

Die Prädiktion der Varianten, die durch WES oder Sanger-Sequenzierung entdeckt wurden, wurden mittels dbNSFP durchgeführt. dbNSFP konnte kostenlos aus dem Internet (<https://sites.google.com/site/jpopgen/dbNSFP>) heruntergeladen und lokal gespeichert werden. Mittels der Befehle in der Eingabeaufforderung wurden die Varianten (nsSNVs und ssSNVs) interpretiert. Die Varianten wurden letztendlich nach der Klassifikation der „ACMG (American College of Medical Genetics and Genomics) Standards and Guidelines“ vorgenommen (Richards et al., 2015).

5.5.7 Transkriptom-Analyse mittels R

Die Daten, die mittels RNA-Sequenzierung erhoben wurden, wurden in der Software R mit Hilfe von edgeR v3.16.5 und limma v3.30.13 für die Transkriptom-Analyse analysiert (Ritchie et al., 2015; Robinson et al., 2010). Die murinen Gen-IDs wurden mithilfe des Ensembl-BiomaRt-R-Paketes in humane Gen-IDs übersetzt (Durinck et al., 2005, 2009). Das Berechnen der Reads von technischen Replikaten erfolgte mit diesem Programm und die Daten wurden mittels der „edgeR’s Trimmed-Mean of M-values, TMM“ Methode und der limma-VomM-Methode normalisiert. Der Zeitpunkt der Dissektion wurde als Variante für das Regressionsmodell verwendet. Die Regressionsanalyse wurde mit Hilfe von limma-lmFit und der eBayes-Methode, sowie der Treat-Methode mit einem Logfoldchange durchgeführt. Die Bilder von Heatmaps wurden mittels des pheatmap-Paketes v1.0.8 generiert (Kolde, 2015).

eBayes

eBayes ist eine Methode, um differentiell exprimierte Gene zu bestimmen, und entspricht sozusagen dem Test für nominale Signifikanz. Der Mittelwert und die Abweichung der Genexpression wurden auf der Grundlage eines empirischen Bayes-Rahmens eingeschätzt und die Signifikanz der Differenzierung nach dem empirischen Bayes bestimmt (Niemi et al., 2015; Smyth, 2004).

Treat (threshold)

„Treat“ ermöglicht es, die Hypothese zu überprüfen, dass der mit differentieller Expression assoziierte P-Wert eines Experimentes größer als ein gegebener Schwellenwert ist (McCarthy and Smyth, 2009).

5.6 Zellenabstammungsanalyse in Mausembryonen

Die Zellenabstammungsanalyse wurde von der Arbeitsgruppe von Prof. Gen Yamada (Abteilung für Entwicklungs-genetik, Medizinische Universität Wakayama, Japan) in den Mausembryonen durchgeführt, um die Einflüsse der zu untersuchenden Kandidatengene bei der Entwicklung eines fehlgebildeten urogenitalen Gewebes zu klären.

5.7 Whole-mount in-situ-Hybridisierung (WISH)

Die Whole-mount in-situ-Hybridisierung (WISH) wurde von der Arbeitsgruppe von Prof. Benjamin Odermatt (Institut für Anatomie, Universität Bonn, Deutschland) nach Standardprotokollen (Thisse and Thisse, 2008) durchgeführt, um die Genexpression im krankheitsrelevanten urorektalen Gewebe in Zebrafischen nachzuweisen. Die Sonden, die speziell an die gezielte mRNA der zu untersuchenden Kandidatengene hybridisierten, wurden speziell für die WISH in den Zebrafischen hergestellt.

6 Ergebnisse

6.1 Genetische Untersuchungen bei Patienten mit Blasenekstrophie-Epispadie-Komplex (BEEK)

6.1.1 Replikationsstudie bei Patienten mit klassischer Blasenekstrophie (KBE)

In unserer vorherigen zweiten GWAS-Studie von 110 KBE-Patienten und 1.177 Kontrollen mit europäischer Herkunft wurde eine genomweit signifikante Region auf Chromosom 5q11.1-q11.2 identifiziert (Draaken et al., 2015). Diese Region beinhaltet das Gen *ISL1*. Viele Marker in dieser Region waren mit KBE assoziiert, darunter waren 84 genomweit signifikante Marker (Draaken et al., 2015). Eine nachfolgende Meta-Analyse von Daten aus unserer ersten GWAS-Studie von 98 KBE Patienten und 526 Kontrollen mit europäischer Abstammung und der zweiten GWAS-Studie implizierte ebenfalls den 5q11.1-q11.2 Locus als Risikolocus für KBE (s. Abb. 13) (Reutter et al., 2014b; Draaken et al., 2015; Zhang et al., 2017a). In dieser Meta-Analyse waren insgesamt 138 Marker in dem Locus 5q11.1-q11.2 genomweit signifikant (s. Abb. 14) (Draaken et al., 2015). Keine anderen Marker erreichten das genomweite Signifikanzniveau.

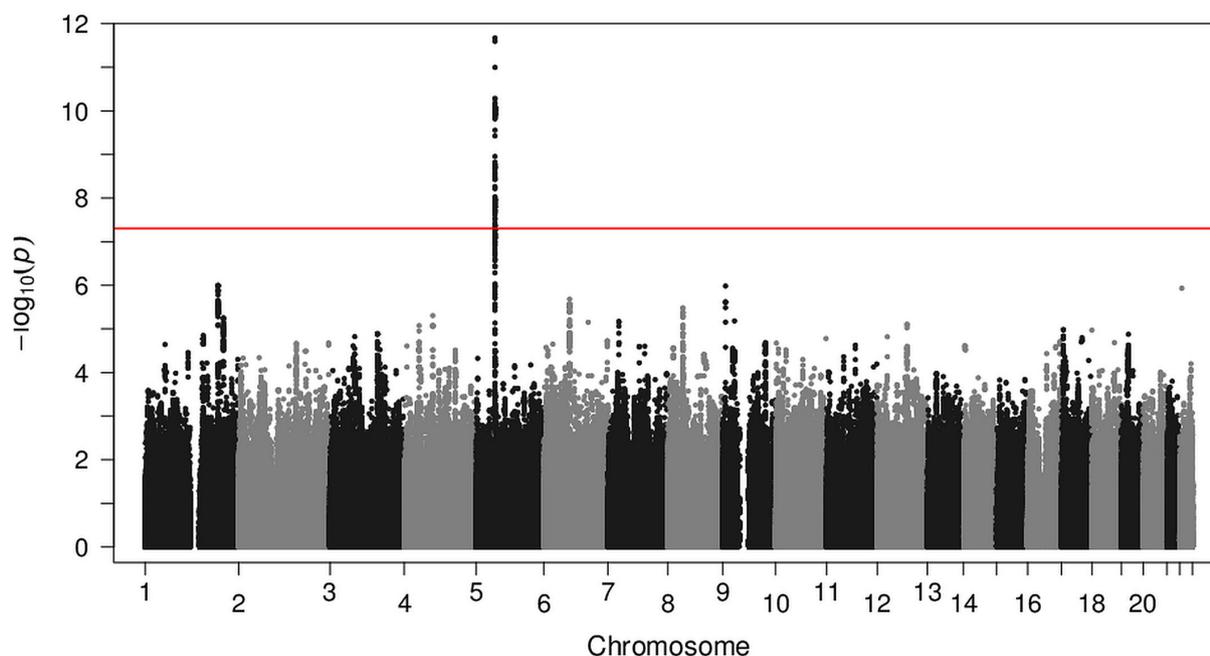


Abbildung 13 Genomweite Assoziationsstudie in Patienten mit klassischer Blasenekstrophie. Die Assoziation der SNPs ist als „ $-\log_{10}(p)$ “ gegen die chromosomale Position grafisch dargestellt. Auf der Y-Achse finden sich die negativen \log_{10} P-Werte der logistischen Regression für SNPs aus der Meta-Analyse, die der Qualitätskontrolle entsprachen. Die Chromosomen sind in wechselnden Farben entlang der X-Achse dargestellt. Das genomweite Signifikanzniveau ist durch eine rote Linie angezeigt (Draaken et al., 2015).

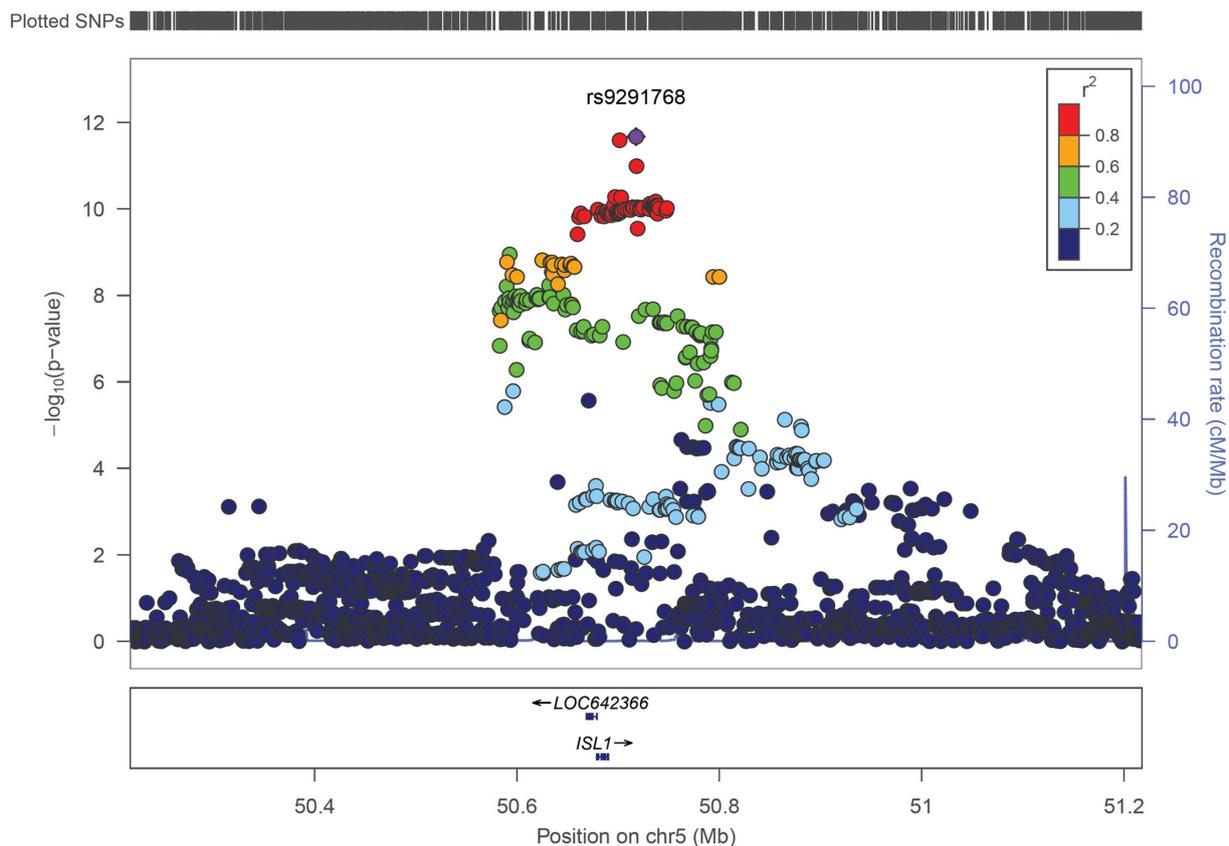


Abbildung 14 Graphische Darstellung der assoziierten Region von *ISL1*. Die Assoziation der SNPs ist als „ $-\log_{10}(p)$ “ gegen die chromosomale Position grafisch dargestellt. Die Y-Achse auf der rechten Seite zeigt die Rekombinationsrate, die aus den EUR-Populationen im 1000-Genomprojekt (März 2012) eingeschätzt wurde. Alle P-Werte (Y-Achse links) stammen aus der Meta-Analyse. Der signifikanteste Marker rs9291768 ist als ein violetter Diamant dargestellt (Draaken et al., 2015).

Um die vielversprechende genomische Region 5q11.1-q11.2 weiter zu verfolgen, wurde eine Replikationsstudie mit 274 KBE-Patienten europäischer Herkunft und 1.365 ethnisch angeglichenen Kontrollen, sowie 110 nordamerikanischen KBE-Eltern-Kind-Trios mit europäischer Abstammung durchgeführt. Hierfür haben wir 12 Marker ausgewählt. Einer davon liegt in der Region von *ISL1* und war in den vergangenen GWAS-Studien und Meta-Analyse genomweit signifikant. Die übrigen 11 Marker waren zwar nominell aber nicht genomweit signifikant und erreichten einen P-Wert von 10^{-05} .

Nach den Qualitätskontrollen für die generierten Genotypdaten wurden sechs KBE-Patienten und 11 ethnisch angeglichenen Kontrollen von der Fall-Kontroll-Studie ausgeschlossen. Daher umfasste der analysierte Datensatz für die Fall-Kontroll-Studie 268 KBE-Patienten bestehend aus Australiern (n = 31), Briten (n = 40), Deutschen (n = 7), Italienern (n = 39), Spaniern (n = 35) und Schweden (n = 116) mit europäischer Herkunft und 1.354 ethnisch angeglichenen Kontrollen. Bei den Kontrollen gab es keine Anzeichen für die Abweichung vom Hardy-

Weinberg Gleichgewicht (engl. Hardy-Weinberg equilibrium, HWE) für alle 12 SNPs (Appendix B).

Von den nordamerikanischen Eltern-Kind-Trios wurden 18 KBE-Trios nach den Qualitätskontrollen des Transmission-Disequilibrium Tests (TDT) von weiteren Analysen ausgeschlossen. Somit enthielt der analysierte Datensatz für den TDT 92 Patient-Eltern-Trios aus Nordamerika jedoch mit europäischer Herkunft.

In der Tabelle 17 werden die Ergebnisse der Assoziationsanalysen und der kombinierten Meta-Analyse der 268 KBE-Patienten, den ethnisch abgestimmten 1.354 Kontrollen und den 92 Eltern-Kind-Trios dargestellt. Eine detailliertere Beschreibung der Ergebnisse vom TDT wird in Tabelle 18 aufgelistet. Von allen getesteten Markern erzielte Marker rs6874700 in der chromosomalen Region 5q11.2, die den *ISL1*-Locus repräsentiert, den niedrigsten P-Wert mit $2,22 \times 10^{-08}$. In der kombinierten Analyse der vorangegangenen Meta-Analyse mit der aktuellen Meta-Analyse zeigte sich ein P-Wert von $9,2 \times 10^{-19}$ für den Marker rs6874700. Das relative Risiko (RR) (95% CI) für rs6874700 betrug demnach 1,93 (95% CI, 1,67-2,23) (Tabelle 19). Der genomweit signifikante Locus konnte somit erfolgreich in den neuen KBE-Kohorten repliziert werden.

Tabelle 17 Ergebnisse der Assoziationsanalysen und der kombinierten Meta-Analyse der KBE Patienten und Kontrollen sowie der Patient-Eltern-Trios (Zhang et al., 2017a).

SNP	Chromosom/ Region	Position	Risiko/ anderes Allel	Meta-analysis Cases (n = 268) vs. controls (n = 1.354) (AUS/GB/DE/ IT/ES/SE) RR (95% CI)	TDT Patient-Eltern Trio (n = 92) (Nord Amerika) RR (95% CI)	Meta-Analyse Patienten (n = 268), Kontrollen (n = 1.354), Patient- Eltern Trios (n = 92) RR (95% CI)	p Werte
rs1475601	1q31.3	194.721.655	G/A	1,11 (0,60–2,05)	0,88 (0,32–2,41)	1,04 (0,61–1,76)	0,884
rs72748303	1q32.2	208.973.633	A/G	1,54 (0,95–2,52)	0,71 (0,23–2,25)	1,37 (0,87–2,15)	0,170
rs79145914	4p14	39.490.706	G/A	1,26 (0,74–2,14)	1,75 (0,51–5,98)	1,33 (0,82–2,15)	0,254
rs7689350	4q13.3	76.207.570	A/C	0,98 (0,70–1,38)	2,13 (0,92–4,92)	1,10 (0,80–1,51)	0,568
rs6874700	5q11.2	50.701.750	A/T	1,80 (1,44–2,25)	1,61 (1,07–2,41)	1,75 (1,44–2,13)	2,2 × 10⁻⁰⁸
rs1514351	6q12	68.694.074	C/T	1,20 (0,96–1,50)	0,91 (0,60–1,39)	1,13 (0,92–1,37)	0,239
rs117622209	6q22.1	118.108.159	C/T	1,12 (0,64–1,96)	1,20 (0,37–3,93)	1,13 (0,69–1,88)	0,623
rs56189125	7p14.3	28.855.348	G/A	1,05 (0,09–11,78)	1,67 (0,61–4,59)	1,56 (0,61–3,96)	0,353
rs57086087	8p11.21	40.301.811	T/C	1,07 (0,80–1,44)	1,42 (0,84–2,39)	1,15 (0,89–1,48)	0,301
rs10119066	9p24.1	7.509.895	T/G	0,99 (0,66–1,50)	1,60 (0,52–4,89)	1,05 (0,71–1,55)	0,806
rs16917077	9p21.1	31.532.212	C/T	1,39 (0,74–2,59)	1,38 (0,55–3,42)	1,38 (0,83–2,32)	0,215
rs1514921	12q21.2	79.481.756	C/T	1,21 (0,89–1,64)	1,35 (0,72–2,53)	1,23 (0,94–1,62)	0,136

AUS: Australiern; GB: Briten; DE: Deutschen; IT: Italienern; ES: Spanien; SE: Schweden; RR: relatives Risiko; CI: Konfidenz Interval; TDT: transmission disequilibrium test.

Tabelle 18 Ergebnisse vom TDT in 92 Patient-Eltern-Trios (Zhang et al., 2017a).

SNP	Chromosom/Region	Position	Risiko-/anderes Allel	T	NT	TDT	P_TDT
rs1475601	1q31.3	194.721.655	G/A	8	7	0,07	7,96E-01
rs72748303	1q32.2	208.973.633	A/G	7	5	0,33	5,64E-01
rs79145914	4p14	39.490.706	G/A	7	4	0,82	3,66E-01
rs7689350	4q13.3	76.207.570	A/C	17	8	3,24	7,19E-02
rs6874700	5q11.2	50.701.750	A/T	61	38	5,34	2,08E-02
rs1514351	6q12	68.694.074	C/T	46	42	0,18	6,70E-01
rs117622209	6q22.1	118.108.159	C/T	6	5	0,09	7,63E-01
rs56189125	7p14.3	28.855.348	G/A	10	6	1,00	3,17E-01
rs57086087	8p11.21	40.301.811	T/C	34	24	1,72	1,89E-01
rs10119066	9p24.1	7.509.895	T/G	8	5	0,69	4,05E-01
rs16917077	9p21.1	31.532.212	C/T	11	8	0,47	4,91E-01
rs1514921	12q21.2	79.481.756	C/T	23	17	0,90	3,43E-01

T: Anzahl der Transmissionen des Risikoallels von heterozygoten Eltern; NT: Anzahl der Nichttransmissionen des Risikoallels von heterozygoten Eltern; P_TDT: P-Wert für TDT.

Tabelle 19 Ergebnisse von unseren vorherigen und aktuellen Meta-Analysen für den Marker rs6874700 (Zhang et al., 2017a).

SNP	Chromosom/Region	Position	Risiko-/anderes Allel	Vorherige Meta-Analyse Draaken et al. 2015 RR (95% CI)	Replikation RR (95% CI)	Vorliegende Meta-Analyse RR (95% CI) p Wert
rs6874700	5q11.2	50.701.750	A/T	2,17 (1,75–2,70)	1,75 (1,44–2,13)	1,93 (1,67–2,23) 9,2 × 10 ⁻¹⁹

CI: Konfidenz Interval; RR: relatives Risiko.

6.1.2 Kandidatengenanalyse in Patienten mit BEEK

Sequenzanalyse für *EFNB1*

In unseren CNV-Studien von 169 BEEK-Patienten wurde eine große vom gesunden Vater vererbte Duplikation mit einer Größe von 6,08Mb in der chromosomalen Region Xq11.1-q13.1 bei einer Patientin mit KBE identifiziert, die darüber hinaus ein persistierendes Foramen ovale und eine beidseitige Inguinalhernie zeigte (von Lowtzow et al., 2016). Diese Duplikation umfasst 43 RefSeq Gene unter ihnen *EFNB1*, welches zuvor mit dem Craniofrontonasalen-Syndrom (MIM #304110) assoziiert wurde und nur im weiblichen Geschlecht auftritt (von Lowtzow et al., 2016). Interessanterweise wurde, das gleichzeitige Auftreten von KE und dem Craniofrontonasalen-Syndrom bereits in zwei unabhängigen Patientinnen beobachtet (Ludwig et al., 2009). Obwohl in der obengenannten Patientin keine KE auftrat, konzentrierte sich die nachfolgende Sequenzanalyse auf alle weiblichen KE Patientinnen in unserer Kohorte (n = 25).

Bei der Sequenzanalyse aller kodierenden Regionen und angrenzenden Spleiß-Stellen von *EFNB1* wurde keine Mutation bei 25 KE Patientinnen identifiziert. In einer Patientin wurde im Exon 5 eine extrem seltene synonyme Variante (rs143341175, p. Ser281=) nachgewiesen. Bei vier Patienten wurde ein häufiger Polymorphismus im 3'-UTR (rs2230423, C/T, MAF 0,1 in der europäischen Population) nachgewiesen (von Lowtzow et al., 2016).

Sequenzanalyse für *ISL1*

In der letzten GWAS-Studie wurde der erste genomweit signifikante Locus für KBE in der chromosomalen Region 5q11.1-q11.2 identifiziert und in der vorliegenden Replikationsstudie wurde der Locus 5q11.1-q11.2 in unterschiedlichen Kohorten repliziert. Außerdem wurde eine starke Expression von *isl1* in der sich entwickelnden Genital- und kloakalen Region durch WISH (engl. whole mount in-situ hybridization) in Mausembryonen gezeigt (Abb. 15) (Draaken et al., 2015). Dies deutet darauf hin, dass mit großer Wahrscheinlichkeit das *ISL1*-Gen ursächlich mit KBE assoziiert ist.

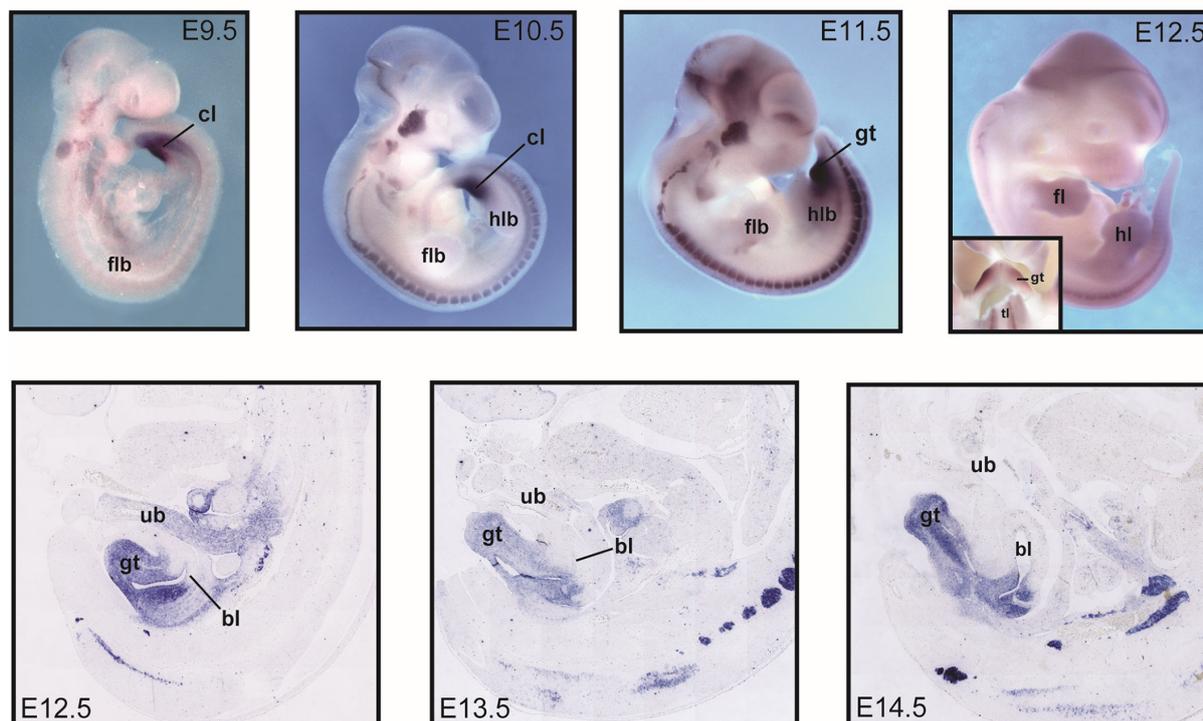
Isl1

Abbildung 15 Expression von *isl1* während der embryonalen Entwicklung der Maus. Die WISH wies eine starke Expression des Gens *isl1* in der entwickelnden Genitalregion (einschließlich der Kloake, der Kloakenmembran und des auftauchenden genitalen Tuberkels) der Wildtyp-Mausembryonen zwischen E9.5 und E12.5 auf. Durch die WISH in den mittleren-sagittalen Paraffin-Schnitten bei späteren embryonalen Stadien (E12.5-E14.5) wurde eine Expression des Gens *isl1* im gesamten genitalen Tuberkel und innerhalb des perikloakalen Mesenchyms und des Urorektalseptums nachgewiesen (Draaken et al., 2015).

Eine erste Sequenzanalyse der kompletten *ISL1*-kodierenden Regionen und Spleißstellen wurde in 207 KBE-Patienten mittels Sanger Sequenzierung durchgeführt. Hierbei zeigten sich weder pathogene „nonsense“ noch pathogene „missense“ Varianten. Wir können jedoch die Möglichkeit nicht ausschließen, dass einige Mutationen übersehen wurden, d.h. Mutationen, die sich in der Promotorregion oder unbekanntem regulatorischen Sequenzen oder den nicht-kodierenden Regionen befinden da die Sequenzanalyse solche Regionen nicht abdeckte. Daher haben wir zusätzlich zwei Intron-Bereiche in der gleichen Kohorte von 207 KBE-Patienten sequenziert, da diese zwei Introns (Intron-1 und 2) im Vergleich mit anderen Spezies sehr konserviert sind (Abb. 16). Auch hier zeigten sich keine neuen Varianten mit Ausnahme eines häufigen SNP rs3811910 (hetero A/G in 43 Patienten, homo G in 6 Patienten).

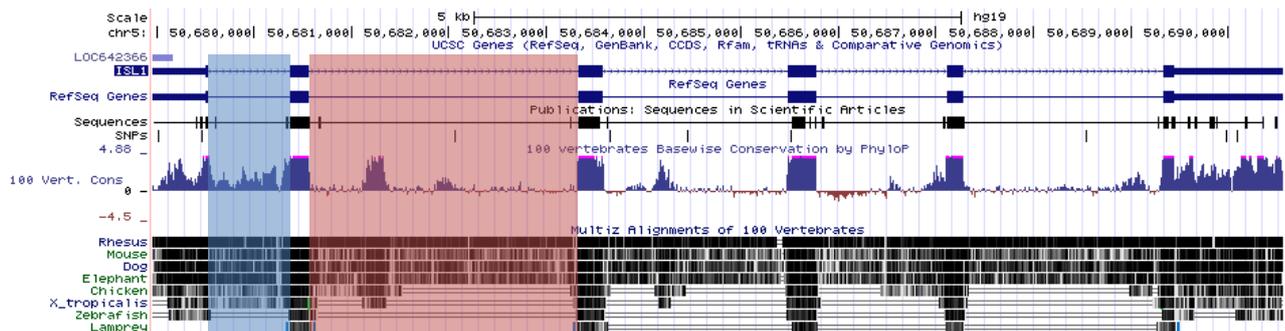


Abbildung 16 Darstellung des Gens *ISL1* im UCSC Genome Browser. Die blaumarkierte Region stellte das Intron-1 dar, Intron-2 ist rot markiert. In dieser Abbildung wurde die Sequenz von *ISL1* (Homo sapiens) und die Konservierung aller Regionen von *ISL1* mit der Sequenz von *isl1* in acht Wirbeltier-Arten verglichen.

Sequenzanalyse für *SLC20A1*

Da der BEEK die Reproduktionsfähigkeit der Betroffenen einschränkt ist davon auszugehen, dass bei einem Teil der Patienten der BEEK auf krankheitsverursachende *de novo* Mutationen zurückzuführen ist. In Übereinstimmung mit dieser Annahme wurden zuvor in der AG Reutter kausale *de novo* Mikrodeletionen und Mikroduplikationen bei BEEK-Patienten identifiziert (Draaken et al. 2014; von Lowtzow et al., 2016). Um nun kleinere *de novo* Mutationen wie z.B. Punktmutationen zu entdecken, wurden acht KE-Patienten mittels WES analysiert (Reutter et al., 2016b). Bei einer Patientin wurde eine nicht-synonyme *de novo* Variante (p.Gly237Arg) in *SLC20A1* identifiziert. Diese *de novo* Mutation wurde mittels achter Prädiktionsalgorithmen in dbNSFP als krankheitsursächlich eingestuft. Außerdem liegt diese Variante nicht in den Exom-Daten von ExAC (<http://exac.broadinstitute.org/gene/ENSG00000144136>) vor. Darüber hinaus charakterisiert ein negativer Residual-Variation-Intolerant-Score (RVIS) (www.genic-intolerance.org) mit einem Wert von -0,49 (22,36%) *SLC20A1* als ein Gen, welches gegenüber neu aufgetretenen genetischen Varianten intolerant ist. Dies deutet darauf hin, dass *SLC20A1* eine mögliche Rolle in wichtigen embryonalen Prozessen und in der Ätiologie des BEEK spielt.

Die bereits existierenden Expressionsdaten in der Maus, zeigen eine starke Expression von *Slc20a1* im „Midgut“ im Theiler-Stadium (TS) 23 (E14.5) (Blake et al., 2017). Während der embryonalen Entwicklung beim Menschen verlängert sich der „Midgut“ von der Spitze der duodenalen Schleife bis zum letzten Drittel des Querkolons. Aufgrund des differenziellen Wachstums von einigen Organen einschließlich den Nieren, der Leber und der kranialen Anteile des Mitteldarms, welche sich schneller verlängern als die kaudalen Anteile, tritt der „Midgut“ durch die Nabelöffnung in die Nabelschnur aus. Diese Herniation beginnt in der

sechsten Woche der menschlichen embryonalen Entwicklung, entsprechend dem kritischen Zeitrahmen der Entstehung der KE beim Menschen. Der gesamte Prozess ist von einer 90° linksdrehenden Rotation begleitet, die an der ventralen Bauchwand des Embryos beobachtet wird (Gearhart, 1998). Infolge des Versagens der Rückverlagerung vom „Midgut“ innerhalb des Nabels zurück in die Bauchhöhle entwickelt sich die menschliche Omphalozele, die zugleich ein phänotypisches Hauptmerkmal der KE und in milder Form der KBE darstellt (Gearhart, 1998). In einer Studie an *Slc20a1*-Knockout-Mäusen konnten Wallingford und Giachelli (2014) zeigen, dass der Prozess der Rückverlagerung des „Midguts“ in den Bauchraum von der Funktion von *Slc20a1* abhängt. Wallingford und Giachelli (2014) konnten zeigen, dass *Slc20a1*-Knockout-Mäuse eine abnorme Endozytose im viszeralem Endoderm des Dottersacks aufweisen, welche dazu führt, dass der Dottersack sich nicht involutieren kann und ein Verschluss der ventralen Bauchwand in Mäuseembryonen ausbleibt entsprechend dem humanen KE-Phänotyp (Wallingford and Giachelli, 2014).

Die Ergebnisse unserer WES Analyse mit der Identifizierung einer *de novo* Mutation in *SLC20A1* in einer Patientin mit KE und der oben genannten Beobachtungen von Wallingford und Giachelli (2014) führten wir eine umfangreiche Re-Sequenzierung von *SLC20A1* bei weiteren 526 Patienten mit BEEK mittels Sanger Sequenzierung durch. Alle kodierenden Regionen und Spleißstellen von *SLC20A1* wurden durch eigens entworfene Primer für *SLC20A1* abgedeckt und insgesamt zwei weitere Varianten in zwei unabhängigen Patienten identifiziert, die bisher nicht in 110.000 öffentlich zugänglichen Kontrollen beschrieben wurden (Tabelle 20). Eine dieser beiden Varianten trat ebenfalls *de novo* auf, die andere wurde von einer betroffenen Mutter vererbt.

Tabelle 20 Darstellung und Klassifizierung der Varianten in *SLC20A1* bei Patienten mit BEEK.

Ext-Code	Mut cDNA	Mut Protein	Väterlich / Mütterlich Vererbt	SIFT	LRT	Mutation Taster	Mutation Assessor	FATHMM	PROVEAN	MetaSVM	MetaLR	fathmm- MKL_coding	CADD Score	Klassifikation
407_501	c.709G>A	G237R	<i>de novo</i>	D	D	D	M	D	D	D	D	D	31	pathogen
188_501	c.893T>C	V298A	<i>de novo</i>	T	N	N	N	D	N	T	T	N	0,001	pathogen
326_501	c.1321A>C	K441Q	mütterlich vererbt	T	D	N	M	D	N	T	D	D	10,24	vermutlich pathogen

D: Krankheits-verursachend; H: hoch funktional; L: nicht-funktional; M: medium funktional; N: neutral; T: tolerabel.

6.1.3 „Lineage Analysis of Celllines“ (Zell-Abstammungsanalysen) in Mausembryonen für *ISL1*

Zell-Abstammungsanalysen ermöglichen es, Kenntnisse über den zellulären Ursprung von Geweben und damit von Erkrankungen/Fehlbildungen zu erlangen. Um den Beitrag von *isl1*-exprimierenden Zellen im sich entwickelnden Urogenitaltrakt im frühen Stadium untersuchen, wurde das *isl1*-mER-Cre-mER-Allel in Mausembryonen verwendet. Dieses Allel stellt eine Tamoxifen-abhängige Cre-Rekombinase-kondierende Sequenz dar. Durch Tamoxifen induziert zeigten sich *isl1*-exprimierende Zellen in dem sich früh entwickelnden Urogenitaltrakt des Mausembryos zum Zeitpunkt E10.5 (Abb. 17A). Zum Zeitpunkt E15.5 verteilten sich die Zellen in der Harnblase der Mäusembryonen (Abb. 17B) (Zhang et al., 2017a).

Um die Beteiligung von *isl1*-exprimierenden Zellen in der Entwicklung der Harnblase weiter zu verfolgen, wurde eine detaillierte Zelllinienanalyse nach Tamoxifen-Behandlung zu verschiedenen Zeitpunkten durchgeführt. Eine zunehmende Anzahl von *isl1*-exprimierenden Zellen bei E7.5 wurde im frühen Stadium der Entwicklung des Urogenitaltraktes und der Region der Harnblase nachgewiesen (Abb. 17C&F). Des Weiteren zeigte sich eine große Anzahl von *isl1*-exprimierenden Zellen in den Regionen des Urogenitaltraktes nach Induktion mit Tamoxifen bei E8.5 und E9.5 (Abb. 17D,E,G,H). Zusätzlich zur Harnblase fanden sich *isl1*-exprimierende Zellen auch in der Region des Genitalhöckers (Abb. 17F-H) (Zhang et al., 2017a).

Um den Beitrag der *isl1*-exprimierenden Zellen zur Entwicklung der kaudalen Körperregion weiter zu untersuchen, wurden *Hoxa3*-IRESCre; *isl1*Flox/Flox-Mutantenembryonen (*isl1* cKO) analysiert. Die *Hoxa3*-Cre-gesteuerten Mäuse besaßen die Cre-Aktivität in den kaudalen Körperregionen (Macatee et al., 2003). Die *isl1*-cKO-Mäuse präsentierten zeigen eine Hypoplasie des unteren Urogenitaltraktes einschließlich der Harnblase und des Genitalhöckers (GT, genital tract) (Abb. 17I-L) (Abb. 17K&L) (Zhang et al., 2017a).

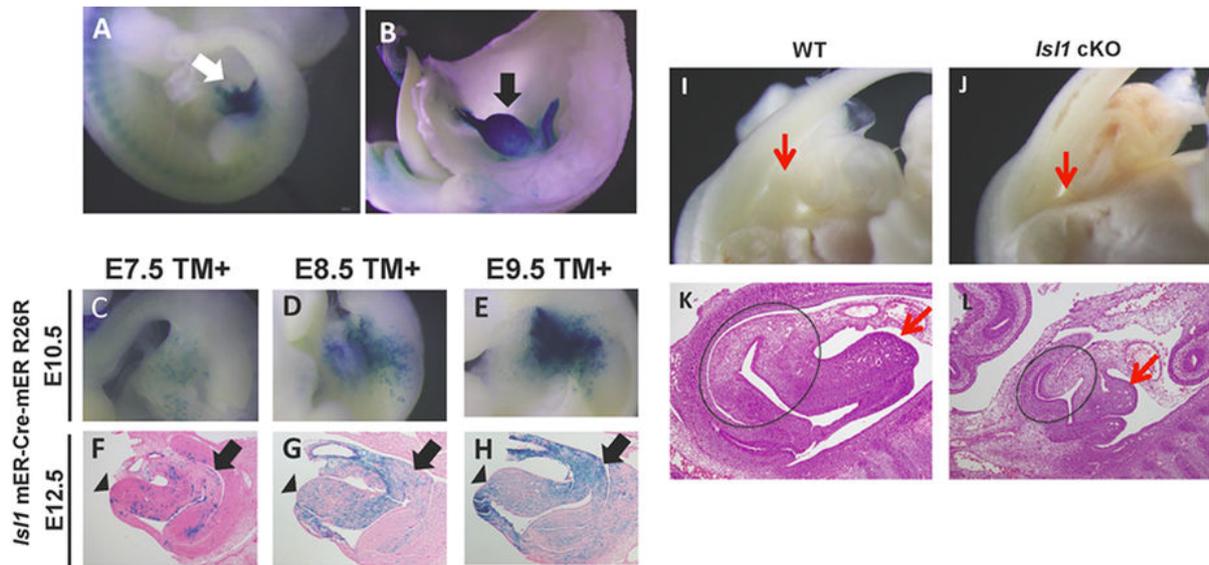


Abbildung 17 Zell-Abstammungsanalyse von *is/1*-exprimierenden Zellen in Mausembryonen. A&B: Zell-Abstammungsanalyse von *is/1*-exprimierenden Zellen bei E10.5 und E15.5 mittels Whole-mount-lacZ-Färbung. A: Beitrag von *is/1*-exprimierenden Zellen in Mausembryonen bei E10.5 (weißer Pfeil); B: Beitrag von *is/1*-exprimierenden Zellen in Mausembryonen bei E15.5 (schwarzer Pfeil); C-H: Zellenabstammungsanalyse von *is/1*-exprimierenden Zellen nach Induktion mit Tamoxifen™ bei E7,5-E9,5; F-H: median-sagittalen Paraffin-Abschnitten der Blasenregion bei E12.5. Die Pfeile und Pfeilspitzen zeigen auf die Harnblase bzw. den unteren Urogenitaltrakt. I-L: Die Abbildung zeigt den hypoplastischen GT und die hypoplastische Harnblase in *is/1* cKO Embryonen (Zhang et al., 2017a).

6.1.4 Whole-mount in-situ Hybridisierung (WISH) in Zebrafisch Larven (zfl)

Im Gegensatz zu Säugetieren besteht der Pronephros aus zwei gegenüberliegenden Glomeruli, zwei pronephrischen Tuben, die in jeweils in einen pronephrischen Harnleiter münden, ehe der Urin dann in die Kloake ausgeschüttet wird. Bereits in den frühen Entwicklungsstadien von zfl zwischen 36-96 Stunden nach Befruchtung (engl. hours post fertilisation, hpf) ist der Pronephros voll funktionsfähig. Im Gegensatz zu ausgewachsenen Fischen bildet das anschließend entwickelte Mesonephros bei Säugetieren die endgültigen Nieren (Marra and Wingert, 2014). In den frühen Entwicklungsstadien von zfl, kurz nach der Gastrulation (Schwanzknospfen-Stadium 10hpf), konnten die ersten renalen Vorläufer in bilateralen Streifen von intermedialem Mesoderm (IM) detektiert werden. Die ersten renalen Vorläufer liegen auf einer der beiden Seiten der embryonalen Mittellinie zwischen dem paraxialen Mesoderm (PM) und dem lateralen Plattenmesoderm (LPM) (Drummond et al., 1998). Ab einem Stadium mit etwa 10 Somiten entwickeln sich diese renalen Vorläufer zu bilateralen Nephronen.

Um die potentielle Beteiligung von *isl1* in der Entwicklung der oben beschriebenen renalen Vorläufer in zfl zu untersuchen, wurde zunächst die Expression von *isl1* an den Seiten während der jeweiligen Zeitpunkte mittels der WISH bei 22 und 60hpf (entsprechend etwa 26 Somiten und des Pec-Fin-Stadiums) in Wildtyp-zfl (wt-zfl) untersucht. Unter Verwendung einer Antisense-in-situ-Sonde, die eine UTR-Region der zfl-*isl1*-cDNA umfasste, konnte die Expression von *isl1* in der Region des sich entwickelndem Pronephros zu allen oben erwähnten Zeitpunkten deutlich dargestellt werden (Zhang et al., 2017a). Bei 48hpf wurde eine deutliche Expression von *isl1* in den anterior gelegenen Regionen des proximalen Tubulus (engl. proximal convoluted tubule, PCT) (Abb. 18) beobachtet. Der zuvor beschriebene *slc20a1a* in-situ-Marker lokalisiert in derselben Region (Wingert et al., 2007; Zhang et al., 2017a). Während es Long-Pec-Stadiums (48hpf) konnte gezeigt werden, dass *isl1* in der Region nahe dem Dottersack exprimiert wird und die Expressionen von *isl1* und *slc20a1a* in der am weitesten anterior gelegenen Region sich entwickelnder Nephronen miteinander überlappt (Abb. 18) (Zhang et al., 2017a).

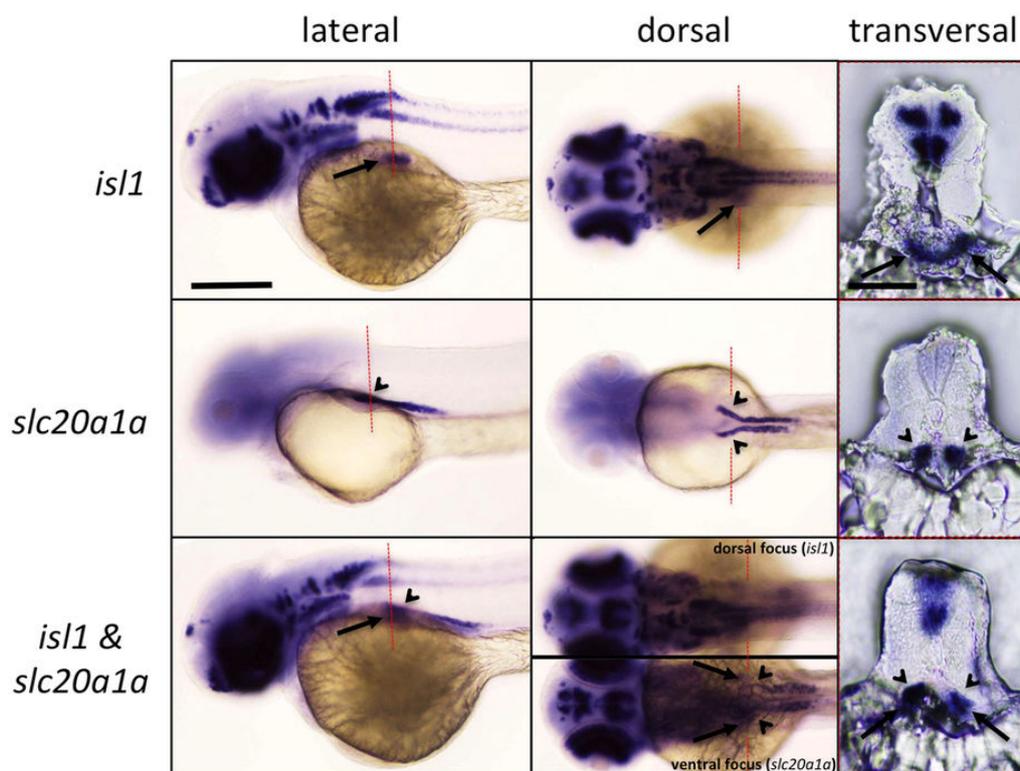


Abbildung 18 WISH gegen *isl1* (Antisense 3'-UTR-Sonde, oben), *slc20a1a* (mitte) und in der Kombination beider Sonden (unten) in Wildtyp-zfl bei 48hpf. Bei dem Long-Pec-Stadium (48hpf) in zfl wurde die Expression von *isl1* in PCT (Pfeile) in lateraler (links) und dorsaler Ansicht (rechts; gleicher Fisch, Skalenstab = 500µm) sowie einem 25µm transversalen Scheibenschnitt des gleichen Fisches (rechte Spalte) dargestellt. Die Ausrichtung der Abschnitte ist mit der roten Linie in den entsprechenden Bildern dargestellt. Die Zfl zeigten eine klare *isl1*-Färbung im Dottersack und die mit *isl1*-gefärbte Region überlappte mit der, mit *slc20a1a* markierten, PCT-Region (schwache Pfeilköpfe) (Zhang et al., 2017a).

6.2 Genetische Untersuchungen bei Patienten mit VATER/VACTERL-Assoziation und VATER/VACTERL-ähnlichen Phänotypen

Unsere vorherigen Studien über die Kopienzahlveränderungen (CNV) lieferten einen Hinweis, dass die kausalen *de novo* CNVs an der Entstehung der VATER/VACTERL-Assoziation und VATER/VACTERL-ähnlichen Phänotypen beteiligt sind. Daher wurde in dieser Arbeit eine SNP-Array-basierte Genotypisierung von 75 Patienten mit VATER/VACTERL-Assoziation und 40 Patienten mit VATER/VACTERL-ähnlichen Phänotypen durchgeführt.

Insgesamt konnten 13 Mikrodeletionen und sieben Mikroduplikationen in 14 Patienten mit VATER/VACTERL-Assoziation und sechs Patienten mit VATER/VACTERL-ähnlichen Phänotypen verifiziert werden (Tabelle 21). Darunter fand sich unter anderem eine *de novo* Mikrodeletion mit einer Größe von 1,39Mb in der chromosomalen Region 17q12 bei einer Patientin mit einer anorektalen Malformation (Vestibuläre Fistel) und Nierenzysten. Außerdem wurden drei weitere Mikrodeletionen in den chromosomalen Regionen 5q23.1, 16q23.3 und 22q11.21 (vermittelt von „low copy repeats“ (LCR) 22s A und D), sowie eine Mikroduplikation in der chromosomalen Region 10q11.21 verifiziert. Da jeweils nur ein Elternteil vorhanden war, konnte für keine der drei Mikrodeletionen (5q23.1, 16q23.3 und 22q11.21) und auch nicht für die Mikroduplikation 10q11.21 ein *de novo* Nachweis geführt werden. Die Mikrodeletion in der chromosomalen Region 22q11.21 wurde bereits zuvor mit VATER/VACTERL-Phänotypen assoziiert und wurde daher als wahrscheinlich ursächlich bei dieser Patientin beurteilt (Brosens et al., 2016; Schramm et al., 2011). Alle übrigen neun Mikrodeletionen und sechs Mikroduplikationen wurden von einem gesunden Elternteil vererbt (Tabelle 21). Drei der identifizierten CNVs (7q35del, 13q31.3del und 13q32.3-q33.1del) überlappten mit zuvor berichteten strukturellen genomischen Aberrationen bei Patienten mit VATER/VACTERL-Assoziation bzw. Patienten mit VATER/VACTERL-ähnlichen Phänotypen (Zhang et al., 2017b).

Tabelle 21 Charakterisierung der chromosomalen Mikroaberrationen in 20 Patienten mit VACTER/VACTERL-Assoziation/-artigen Phänotypen (Zhang et al., 2017b).

Patient	Geschlecht	Phänotyp	Typ der CNV	Chromosom	Maximale Betroffene Region	Größe (bp)	Väterlich / Mütterlich Vererbt	ClinVar Accession Numm
1	weiblich	V, A, C	Deletion	1q44	Chr1:245.900.039-246.475.061	575.023	väterlich vererbt	SCV000537140
2	weiblich	A, C, TE, L	Duplikation	2q21.1q21.2	Chr2:132.140.614-132.312.325	171.711	väterlich vererbt	SCV000537141
3	weiblich	A, R, L	Deletion	4q13.3	Chr4:75.720.151-75.883.784	163.634	mütterlich vererbt	SCV000537142
4	männlich	A, C, TE	Deletion	5q22.1q22.2	Chr5:111.241.351-111.390.356	149.006	väterlich vererbt	SCV000537143
5	weiblich	A, TE	Deletion	5q23.1	Chr5:118.890.915-119.082.631	191.717	?	SCV000537144
6	weiblich	A, R	Duplikation	7q21.2	Chr7:15.424.383-15.534.832	110.449	väterlich vererbt	SCV000537145
7	weiblich	A, TE	Deletion	7q35	Chr7:147.516.080-147.803.861	287.782	väterlich vererbt	SCV000537146
8	männlich	A, R	Deletion	9q22.32	Chr9:97.290.466-97.419.146	128.681	väterlich vererbt	SCV000537147
9	weiblich	V, A, C, R	Duplikation	10q11.21	Chr10:43.796.180-44.037.408	241.228	?	SCV000537148
10	männlich	V, A, R	Duplikation	10q25.1	Chr10:107.722.094-109.081.424	1.359.331	mütterlich vererbt	SCV000537149
11	männlich	V, A, C, R	Deletion	12q23.1	Chr12:100.367.541-100.489.446	121.906	väterlich vererbt	SCV000537150
12	weiblich	A, C, R	Duplikation	13q14.11	Chr13:43.505.151-43.762.719	257.568	väterlich vererbt	SCV000537151
13	männlich	V, A, TE	Deletion	13q31.3	Chr13:93.377.562-93.606.814	229.253	mütterlich vererbt	SCV000537152
14	männlich	V, A, TE, R	Deletion	13q32.3q33.1	Chr13:101.605.129-101.748.121	142.993	väterlich vererbt	SCV000537153
15	weiblich	V, A, R	Deletion	15q25.2q25.3	Chr15:84.931.022-85.728.834	797.813	mütterlich vererbt	SCV000537154
16	männlich	V, A, R	Deletion	16q23.3	Chr16:83.414.018-83.518.410	104.393	?	SCV000537155
17	männlich	A, C, TE, R	Duplikation	17p13.3	Chr17:1.007.540-1.254.875	247.335	mütterlich vererbt	SCV000537156
18	weiblich	A, R	Deletion*	17q12	Chr17:34.815.551-36.208.392	1.392.842	<i>de novo</i>	SCV000537157
19	männlich	A, C, R	Duplikation	17q25.3	Chr17:80.881.583-81.060.040	178.457	mütterlich vererbt	SCV000537158
20	weiblich	A, R	Deletion*	22q11.21	Chr22:18.915.347-21.463.730	2.548.384	?	SCV000537159

* krankheitsursächliche CNV.

6.3 Genetische Untersuchungen bei Patienten mit Ösophagusatresien mit und ohne tracheo-ösophagealer Fistel (EA/TEF)

6.3.1 Whole-Exome Sequencing (WES) bei Patienten mit Ösophagusatresien mit und ohne tracheo-ösophagealer Fistel (EA/TEF)

Mittels Array-basierter CNV-Analyse wurden im Vorfeld bereits strukturell ursächliche chromosomale (*de novo*) Aberrationen bei Patienten mit EA/TEF identifiziert (Brosens et al., 2016). Neben großen chromosomalen *de novo* Aberrationen können auch kleine genetische Varianten eine Rolle bei der Entstehung von EA/TEF spielen. Daher wurde in der vorliegenden Arbeit das Whole-Exome Sequencing (WES) in 31 Eltern-Kind-Trios durchgeführt, um kleine genetische *de novo* Varianten in Patienten mit EA/TEF zu identifizieren.

Insgesamt wurden 26 *de novo* Mutationen in 19 unabhängigen Patienten durch WES entdeckt. 24 *de novo* Mutationen konnten in 18 unabhängigen Patienten mithilfe der Sanger Sequenzierung bestätigt werden (Tabelle 22). Mittels des Algorithmus dbNSFP wurden alle 24 Varianten charakterisiert (Tabelle 23).

6.3.2 Transkriptom-Analyse

Um zu beurteilen, ob die Gene, die durch WES identifiziert wurden, direkt an der embryonalen Entwicklung des Ösophagus beteiligt sind, wurden Transkriptom-Analysen in von Mäusembryonen gewonnenem Gewebe durchgeführt. Die Gewebeproben wurden im Bereich des sich entwickelnden Ösophagus bzw. der sich abspaltenden Trachea zu den Zeitpunkten E8.5 (TS13) und E12.5 (TS21) entnommen. Die Auswertung der Transkriptomdaten erfolgte durch zwei Methoden „eBayes“ und „TREAT“ aus dem Paket „limma“ und die Methode „heatmap“ wurde für Visualisierung der Expressionsdaten benutzt. Nach der Analyse mittels eBayes zeigte sich, dass 13 Gene signifikant differentiell exprimiert waren (rotmarkiert, Tabelle 24). Bei TREAT spielt zusätzlich der „Fold-Change (FC)“ eine Rolle. Ist dieser kleiner als der gesetzte Wert von 1,7 (interner Standardwert), ist die Expression eines Gens zwischen beiden Zeitpunkten nicht differentiell. Hier zeigten nur vier von 20 Genen eine differentielle Expression (rotmarkiert, Tabelle 25). Unter Verwendung der Voom-

normalisierten Genexpression (log₂-Skala) zeigten sieben Gene zwischen E8.5 und E12.5 eine differentielle Expression (Abb. 19).

Tabelle 22 Liste der 24 bestätigten de novo Varianten durch WES in 31 Patient-Eltern-Trios

Ext-Code	HGNC	Refseq	dbSNP-ID, gnomAD	MutCDNA	MutProt	Mm	Gg	Dr	Xt
4_501	<i>EEF1D</i>	NM_032378.4		c.874C>T	p.R292*	K	E	-	K
4_501	<i>CELSR1</i>	NM_014246.1	dbSNP:rs370934549, gnomAD:3/276978	c.4357G>A	p.V1453I	V	I	-	I
21_501	<i>HPS3</i>	NM_032383.3	dbSNP:rs121908316, gnomAD:10/277112	c.1189C>T	p.R397W	H	R	R	R
27_501	<i>PIGC</i>	NM_153747.1		c.716C>T	p.A239V	A	-	G	A
35_501	<i>NFX1</i>	NM_002504.4		c.1723G>A	p.V575M	V	-	-	V
36_501	<i>ZFHX3</i>	NM_006885.3		c.1601C>G	p.P534R	P	P	P	N
750_501	<i>ZFHX3</i>	NM_006885.3	n.f.i. dbSNP, ExAC Browser a.c. 3/120598	c.6377C>T	p.A2126V	A	T	T	A
41_501	<i>MTA3</i>	NM_020744.2		c.393C>A	p.F131L	F	-	F	-
46_501	<i>FANCB</i>	NM_152633.2		c.782G>A	p.R261Q	R	Q	S	-
46_501	<i>PLEC</i>	NM_201379.1	dbSNP:rs376311731, gnomAD:17/269188	c.6704G>A	p.R2235H	R	R	K	R
63_501	<i>PIIP5K2</i>	NM_015216.2	dbSNP:rs781849805, gnomAD:2/242778	c.686G>A	p.R229Q	R	R	R	R
88_501	<i>CLP1</i>	NM_006831.2	gnomAD:1/246264	c.814C>A	p.H272N	H	H	H	H
88_501	<i>GPR133</i>	NM_198827.3	dbSNP:rs572513767, gnomAD:6/276842	c.1033G>A	p.A345T	A	-	-	-
88_501	<i>SLC5A2</i>	NM_003041.3		c.644T>C	p.L215P	L	-	L	L
90_501	<i>KIAA0556</i>	NM_015202.2		c.3730C>T	p.H1244Y	H	H	H	H
141_501	<i>STAB1</i>	NM_015136.2	dbSNP:rs749427730, gnomAD:7/243124	c.6145C>T	p.R2049C	R	-	S	-
154_501	<i>GGT6</i>	NM_153338.2	dbSNP:rs777462893, gnomAD:1/241418	c.1045A>G	p.S349G	S	-	-	-
167_501	<i>CHD7</i>	NM_017780.3		c.4187C>G	p.A1396G	A	A	A	-
167_501	<i>TPP2</i>	NM_003291.2	dbSNP:rs73578896, gnomAD:303/275928	c.1534G>A	p.V512I	I	I	-	M
172_501	<i>NPR2</i>	NM_003995.3		c.952C>G	p.R318G	R	K	-	T
174_501	<i>UBA3</i>	NM_198195.1		c.1088C>T	p.S363F	S	S	T	P
181_501	<i>TANC2</i>	NM_025185.3		c.2357C>T	p.P786L	P	-	P	P
288_501	<i>TRPS1</i>	NM_014112.2		c.1630C>T	p.R544*	R	-	R	R
288_501	<i>APOL2</i>	NM_145637.1		c.319G>C	p.E107Q	D	-	-	-

Mm: *Mus musculus*; Gg: *Gallus gallus*; Dr: *Danio rerio*; Xt: *Xenopus tropicalis*.

Tabelle 23 Charakterisierung der de novo Varianten

Chr	Position	Ref-Allel	Alt-Allel	Gen	SIFT	LRT	Mutation Taster	Mutation Assessor	FATHMM	PROVEAN	MetaSVM	MetaLR	fathmm-MKL_coding	CADD Score	Klassifikation
1	172411047	G	A	<i>PIGC</i>	T	D	D	M	T	N	T	T	D	11,61	vermutlich pathogen
2	42871278	C	A	<i>MTA3</i>	D	D	D	H	D	D	D	D	D	26	vermutlich pathogen
3	52555698	C	T	<i>STAB1</i>	T	N	N	M	T	D	T	T	N	24,1	unklar
3	69105388	G	A	<i>UBA3</i>	D	D	D	L	T	D	T	T	D	27,8	vermutlich pathogen
3	148868411	C	T	<i>HPS3</i>	D	D	A	M	T	D	D	T	D	35	unklar
5	102483678	G	A	<i>PPIP5K2</i>	D	D	D	M	T	D	T	T	D	34	unklar
8	61750228	C	G	<i>CHD7</i>	D	D	D	H	T	D	D	D	D	33	pathogen
8	116616566	G	A	<i>TRPS1</i>	.	D	A	D	36	pathogen
8	144671378	G	A	<i>EEF1D</i>	.	N	A	N	28,5	vermutlich pathogen
8	144997327	C	T	<i>PLEC</i>	D	U	D	N	T	N	T	T	D	26,5	unklar
9	33318942	G	A	<i>NFX1</i>	D	N	N	L	T	N	T	T	D	20,8	vermutlich pathogen
9	35799693	C	G	<i>NPR2</i>	T	N	D	L	D	D	T	T	D	22,2	vermutlich pathogen
11	57428444	C	A	<i>CLP1</i>	T	D	D	L	T	N	T	T	D	17,39	unklar
12	131487736	G	A	<i>GPR133</i>	T	N	N	N	T	N	T	T	N	0,016	unklar
13	103288598	G	A	<i>TPP2</i>	T	N	N	N	T	N	T	T	D	14,76	unklar
16	27772832	C	T	<i>KIAA0556</i>	D	N	N	L	T	N	T	T	N	1,985	vermutlich pathogen
16	31498710	T	C	<i>SLC5A2</i>	D	D	D	H	D	D	D	D	D	27,6	vermutlich pathogen
16	72830204	G	A	<i>ZFH3</i>	D	N	D	L	T	N	T	T	D	22,3	vermutlich pathogen
16	72992444	G	C	<i>ZFH3</i>	D	D	D	L	T	D	D	D	D	19,2	vermutlich pathogen
17	4461651	T	C	<i>GGT6</i>	D	N	N	N	T	D	T	T	N	5,968	unklar
17	61432748	C	T	<i>TANC2</i>	T	D	D	L	T	D	T	T	D	22,7	vermutlich pathogen
22	36624145	C	G	<i>APOL2</i>	T	N	N	N	T	N	T	T	N	0,004	vermutlich pathogen
22	46835135	C	T	<i>CELSR1</i>	T	U	N	N	T	N	T	T	D	15,33	unklar
X	14882851	C	T	<i>FANCB</i>	T	N	N	N	T	N	T	T	N	7,239	vermutlich benigne

A: sicher krankheitsverursachend; D: sehr wahrscheinlich krankheitsverursachend; H: hoch funktional; L: nicht-funktional; M: medium funktional; N: neutral; T: tolerabel.

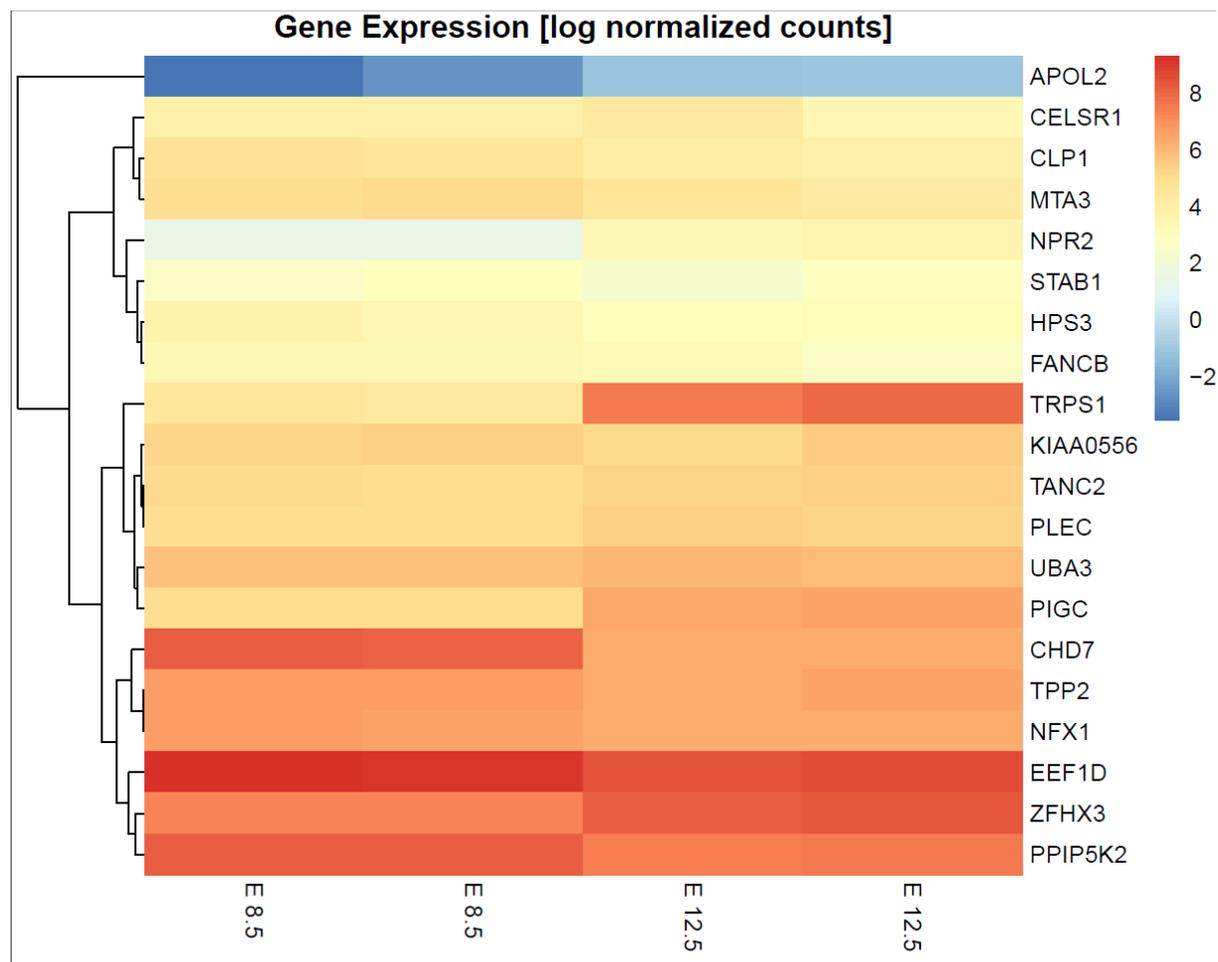
Tabelle 24 Ergebnisse ermittelter Kandidatengene mittels der eBayes-Methode

Murine Gene ID	Human Gene ID	HGNC Symbol	Chromosom	logFC	AveExpr	t	P-Wert	adj. P-Wert	B
ENSMUSG00000041235	ENSG00000171316	<i>CHD7</i>	8	-1,92	7,24	-22,76	1,19E-07	1,50E-05	8,57
ENSMUSG00000038679	ENSG00000104447	<i>TRPS1</i>	8	3,34	6,06	20,38	2,48E-07	2,17E-05	7,90
ENSMUSG00000038872	ENSG00000140836	<i>ZFHX3</i>	16	0,96	7,81	12,97	4,94E-06	1,03E-04	4,46
ENSMUSG00000026698	ENSG00000135845	<i>PIGC</i>	1	1,45	5,73	12,45	6,45E-06	1,19E-04	4,37
ENSMUSG00000028469	ENSG00000159899	<i>NPR2</i>	9	2,11	2,43	8,38	8,16E-05	5,93E-04	2,14
ENSMUSG00000040648	ENSG00000145725	<i>PPIP5K2</i>	5	-0,69	7,84	-7,91	1,17E-04	7,49E-04	0,91
ENSMUSG00000055762	ENSG00000104529	<i>EEF1D</i>	8	-0,67	8,91	-7,40	1,76E-04	9,88E-04	0,37
ENSMUSG00000027079	ENSG00000172409	<i>CLP1</i>	11	-0,66	4,32	-4,71	2,39E-03	6,45E-03	-2,01
ENSMUSG00000055817	ENSG00000057935	<i>MTA3</i>	2	-0,66	4,74	-4,21	4,30E-03	1,02E-02	-2,74
ENSMUSG00000028423	ENSG00000086102	<i>NFX1</i>	9	-0,36	6,45	-4,21	4,27E-03	1,02E-02	-2,98
ENSMUSG00000050014	ENSG00000128335	<i>APOL2</i>	22	1,97	-2,11	3,02	2,01E-02	3,59E-02	-2,98
ENSMUSG00000027615	ENSG00000163755	<i>HPS3</i>	3	-0,60	3,24	-3,60	9,19E-03	1,90E-02	-3,23
ENSMUSG00000022565	ENSG00000178209	<i>PLEC</i>	8	0,31	5,16	2,84	2,59E-02	4,44E-02	-4,73
ENSMUSG00000053580	ENSG00000170921	<i>TANC2</i>	17	0,31	5,19	2,72	3,05E-02	5,02E-02	-4,91
ENSMUSG00000047757	ENSG00000181544	<i>FANCB</i>	X	-0,43	3,12	-2,14	7,08E-02	1,03E-01	-5,33
ENSMUSG00000041763	ENSG00000134900	<i>TPP2</i>	13	-0,25	6,53	-2,34	5,29E-02	8,02E-02	-5,68
ENSMUSG00000030061	ENSG00000144744	<i>UBA3</i>	3	0,18	5,86	1,77	1,21E-01	1,64E-01	-6,41
ENSMUSG00000042286	ENSG00000010327	<i>STAB1</i>	3	-0,29	2,68	-1,23	2,60E-01	3,16E-01	-6,42
ENSMUSG00000016028	ENSG00000075275	<i>CELSR1</i>	22	0,01	3,89	0,03	9,78E-01	9,82E-01	-7,52
ENSMUSG00000032743	ENSG00000047578	<i>KIAA0556</i>	16	0,02	5,30	0,12	9,10E-01	9,26E-01	-7,80

Tabelle 25 Ergebnisse von Kandidatengenens mittels TREAT mit einem „Fold Change Cutoff“ von $\log_2(1,7)$

Murine Gene ID	Human Gene ID	HGNC Symbol	Chromosom	logFC	AveExpr	t	P-Wert	adj. P-Wert
ENSMUSG00000038679	ENSG00000104447	TRPS1	8	3,34	6,06	15,71	7,32E-07	7,66E-05
ENSMUSG00000041235	ENSG00000171316	CHD7	8	-1,92	7,24	-13,69	1,74E-06	1,27E-04
ENSMUSG00000026698	ENSG00000135845	PIGC	1	1,45	5,73	5,87	3,50E-04	4,43E-03
ENSMUSG00000028469	ENSG00000159899	NPR2	9	2,11	2,43	5,34	6,05E-04	6,63E-03
ENSMUSG00000038872	ENSG00000140836	ZFHX3	16	0,96	7,81	2,62	1,76E-02	8,30E-02
ENSMUSG00000050014	ENSG00000128335	APOL2	22	1,97	-2,11	1,85	5,62E-02	2,11E-01
ENSMUSG00000055817	ENSG00000057935	MTA3	2	-0,66	4,74	0,00	7,41E-01	1,00E+00
ENSMUSG00000027079	ENSG00000172409	CLP1	11	-0,66	4,32	0,00	7,57E-01	1,00E+00
ENSMUSG00000040648	ENSG00000145725	PIIP5K2	5	-0,69	7,84	0,00	8,07E-01	1,00E+00
ENSMUSG00000027615	ENSG00000163755	HPS3	3	-0,60	3,24	0,00	8,19E-01	1,00E+00
ENSMUSG00000055762	ENSG00000104529	EEF1D	8	-0,67	8,91	0,00	8,43E-01	1,00E+00
ENSMUSG00000047757	ENSG00000181544	FANCB	X	-0,43	3,12	0,00	9,30E-01	1,00E+00
ENSMUSG00000042286	ENSG00000010327	STAB1	3	-0,29	2,68	0,00	9,62E-01	1,00E+00
ENSMUSG00000053580	ENSG00000170921	TANC2	17	0,31	5,19	0,00	9,98E-01	1,00E+00
ENSMUSG00000022565	ENSG00000178209	PLEC	8	0,31	5,16	0,00	9,98E-01	1,00E+00
ENSMUSG00000016028	ENSG00000075275	CELSR1	22	0,01	3,89	0,00	9,99E-01	1,00E+00
ENSMUSG00000028423	ENSG00000086102	NFX1	9	-0,36	6,45	0,00	9,99E-01	1,00E+00
ENSMUSG00000041763	ENSG00000134900	TPP2	13	-0,25	6,53	0,00	9,99E-01	1,00E+00
ENSMUSG00000030061	ENSG00000144744	UBA3	3	0,18	5,86	0,00	1,00E+00	1,00E+00
ENSMUSG00000032743	ENSG00000047578	KIAA0556	16	0,02	5,30	0,00	1,00E+00	1,00E+00

Abbildung 19 Heatmap der Expression von Kandidatengenen mittels der Voom-Methode



6.3.3 Kandidatengenanalyse für *EEF1D*

Durch WES wurde eine „nonsense“ Mutation im Gen für den Elongationsfaktor 1-Delta (*EEF1D*) identifiziert. Nach in-silico Prädiktionsalgorithmen in dbNSFP wurde diese Mutation (c.874C>T, p.Arg292*) als „automatic disease causing“ von MutationTaster und als „neutral“ von LRT und fathmm-MKL_coding beurteilt. Von den anderen sechs Prädiktionsalgorithmen gab es keine Angabe für diese Nonsense-Mutation, da diese nur nicht-synonyme Mutationen prädiktieren können. Der Analyse von eBayes- und Voom-Methoden nach ist das Gen *Eef1d* bei E8.5 und E12.5 während der Entwicklung des Ösophagus der Maus zu beiden Zeitpunkten exprimiert. Aufgrund der Ergebnisse der WES und der Transkriptom-Analyse priorisierten wir *EEF1D* als das Kandidatengen und führten eine Re-Sequenzierung in 192 EA/TEF-Patienten durch.

Die Sequenzanalyse aller kodierenden Regionen und angrenzenden Spleiß-Stellen von *EEF1D* (NM_001130053) zeigte keine Mutation. Neben einem häufigen heterozygoten exonischen SNP (rs1062391; MAF > 1%), wurden 11 frequente intronische SNPs (MAF > 1%) und 2 seltene heterozygote exonische SNPs (rs780292002 (MAF = 0,0004132), rs78886866 (MAF = 0,001) bei insgesamt 192 EA/TEF-Patienten identifiziert.

7 Diskussion

7.1 Blasenektrophie-Epispadie-Komplex (BEEK)

7.1.1 Replikationsstudie bei Patienten mit BEEK und Analyse des *ISL1*-Locus

Um die vielversprechendsten genomischen Regionen, die auf den Ergebnissen unserer vorherigen Meta-Analyse in insgesamt 208 KBE Patienten basierten, weiter zu verfolgen, konnten in dieser Arbeit die Assoziation mit der *ISL1*-Region und dem Marker rs6874700 ($p = 2,22 \times 10^{-08}$) repliziert werden. In einer Meta-Analyse der hier erhobenen Daten und der zuvor durchgeführten GWAS ergaben sich für den Marker rs6874700 ein kombinierter P-Wert von $9,2 \times 10^{-19}$ mit einem relativen Risiko (RR) (95% CI) von 1,93 (95% CI, 1,67-2,23) (Zhang et al., 2017a).

Das Protein *isl1* ist ein Transkriptionsfaktor, der LIM-Domänen enthält und eine wesentliche Rolle in der Regulierung von Zell-Differenzierungsprozessen in der Organogenese spielt. Zuvor war bekannt, dass es eine essentielle Rolle in der Entwicklung des Herzens, der Nieren, von Neuronen und Gliedmaßen spielt (Biemar et al., 2001; Cai et al., 2003). Allerdings fehlten bislang Untersuchungen, inwieweit *isl1* eine Rolle in der Entwicklung der ableitenden Harnwege spielt. Die Kloake ist eine der essentiellen embryonalen Strukturen für die Entwicklung des externen Genitale (Kaku et al., 2013). Die Epithelien der Kloake und das perikloakale Mesenchym (engl. PCM) und die Interaktionen des PCM spielen eine wichtige Rolle in der kaudalen Entwicklung. Den vorderen Anteil des PCMs wird mit aPCM (engl. anterior part of the PCM, aPCM) bezeichnet und befindet sich im oberen Teil der Kloakenregion (Cai et al., 2003; Haraguchi et al., 2007). Die von unserem Kooperationspartner durchgeführten Abstammungsanalysen zu *isl1*-exprimierenden Zellen in Mäuseembryonen und *isl1*-cKO-Mausanalysen zeigen, dass das aPCM zur embryonalen Entwicklung der Harnblase, des Rektums und des externen Genitale beiträgt. Wie bereits dargestellt haben KBE-Patienten neben der Fehlbildung der Harnblase und des äußeren Genitale zahlreiche Anomalien des menschlichen CAKUT-Spektrums z.B. vesikoureteraler Reflux oder eine Obstruktion des Ureterabgangs vom Nierenbecken (engl. ureteropelvic junction obstruction, UPJ). Da in den Abstammungsanalysen der *isl1*-exprimierenden Zellen sich fast alle Strukturen des sich entwickelnden Harntraktes anfärben, spiegelt dies das gesamte Spektrum der bei KBE-Patienten betroffenen Organesysteme wider. In den aktuellen Studien konnte gezeigt

werden, dass *isl1* für die kaudale Körperbildung einschließlich Harnblase und GT erforderlich ist. Die genaue Abstammungsanalyse der *isl1*-exprimierenden Zellen deutet darauf hin, dass Zellen des aPCM für die Entwicklung der Harnblase und der externen Genitalien bei ca. E7.5 spezifisch sind. Entsprechend könnten genetische Veränderungen der humanen *ISL1*-Region bzw. genetische Veränderungen der kodierenden *ISL1*-Abschnitte oder *ISL1*-regulierender Elemente zu CAKUT-Fehlbildungen beim Menschen führen. Die epithelial-mesenchymale Interaktion zwischen Epithel und PCM ist für die Entstehung des kaudalen Embryos notwendig. Während der kaudalen embryonalen Entwicklung sind viele Signalwege wie die des „Bone-morphogenetischen Proteins (Bmp)“ und des „Sonic Hedgehog-Signalwegs (Shh)“ im aPCM aktiviert. Defekte von Bmp-Signalen führen zu Körperwandverschlussdefekten (Sun et al., 2007; Uchimura et al., 2009). Ein weiterer Wachstumsfaktor, Shh, ist in den Kaudalepithelien im frühen Stadium exprimiert und beeinflusst die Entwicklung des aPCM (Kaku et al., 2013). Es wurde auch gezeigt, dass Shh-Signale die Bmp-Signale modulieren (Haraguchi et al., 2001). Daher deuten die aktuellen Analysen darauf hin, dass die Signale einschließlich Shh und Bmp möglicherweise die *isl1*-Zelllinien des aPCM im frühen embryonalen Stadium wie E7.5 beeinflussen. Weitere Studien sind erforderlich, um die molekularen Mechanismen der Entwicklung der Harnblasen und der Bauchwand zu verstehen (Zhang et al., 2017a).

Mit Hilfe der WISH-Analyse in frühen Entwicklungsstadien von zfl (22-60hpf) konnte die Expression von *isl1* in der Nähe der proximalen Region des sich entwickelnden Pronephros zu allen Untersuchungszeitpunkten nachgewiesen werden. Später während der Entwicklung von zfl fusionieren die beiden Nephrene mit dem kaudalen Ende des Darms und inserieren etwa 96hpf an der Kloake (Pyati et al., 2006). Der Pronephros entspricht der menschlichen Niere und seiner Funktion und filtert das Blut der zfl mit Bildung von Urin, der über die Kloake ausgeschieden wird (Drummond et al., 1998). Die Expression von *isl1* in den proximalen Regionen des sich entwickelnden Harntraktes der Säugetiere und der äquivalenten Strukturen in zfl ergeben zusätzliche Hinweise darauf, dass *ISL1* eine wichtige Rolle in der Entwicklung des Harntraktes bei verschiedenen Arten spielt (Zhang et al., 2017a).

Leider konnte in unserer Einzelmarker Replikationsanalyse keine der anderen 11 Regionen unserer vorangegangenen GWAS repliziert werden. Ein Grund, warum diese 11 Regionen sich in der aktuellen Analyse nicht replizieren ließen könnte darin liegen, dass die Effekte dieser

Regionen in den aktuellen Replikationskohorten geringer sind, als in der Kohorte, die initial in unserer GWAS untersucht wurde. Alternativ könnten diese Regionen in unserer vorherigen Meta-Analyse falsch positive Signale gewesen sein, die nicht für die Ausbildung der KBE relevant sind oder nur für die Ausbildung der KBE bei Patienten mitteleuropäischer Herkunft eine Rolle spielen. Am wahrscheinlichsten ist aber, dass unsere vorherige Meta-Analyse die tatsächlich relevanten Regionen außer der *ISL1*-Region nicht identifiziert hat bzw. diese Regionen einen P-Wert erzielt haben, der größer ist als 10^{-05} . Damit wären diese Regionen in der aktuellen Analyse nicht berücksichtigt. Um daher zusätzliche Marker bzw. Regionen zu identifizieren, die neben *ISL1* an der Entwicklung der menschlichen KBE beteiligt sein könnten, müssen wir genomweite Genotypisierungsanalysen der jetzt mittels Einzelmarkeranalyse untersuchten Kohorten durchführen (Zhang et al., 2017a).

Innerhalb unserer replizierten Regionen mit genomweiter Signifikanz und einem P-Wert von $1,14 \times 10^{-10}$ in unserer vorherigen Meta-Analyse erreicht der Marker rs2303751, der im Exon-4 des Gens *ISL1* vorliegt, einen Wert von von 2b (1a-6) in seiner funktionellen Beurteilung durch die „RegulomeDB“ (<http://regulomedb.org>). Dies deutet darauf hin, dass rs2303751 funktionelle Eigenschaften besitzt. Laut „HaploReg v4.1“ verändert der Marker rs2303751 17 DNA-bindende Motive, einschließlich der Veränderung des PWM (engl. position weight matrix) für *SMAD4*, welches als ein essentieller Transkriptionsfaktor des TGF-beta-Signalwegs agiert (Yan et al., 2014). Darüber hinaus bindet *EZH2* (engl. enhancer of zeste 2) an die genomische Region, die rs2303751 umfasst, *EZH2* ist ein Hauptbestandteil des Genom-regulierenden *PRC2* (engl. poly repressive complex 2) und ein therapeutisches Target bei menschlichen Blasenkrebs (Martínez-Fernández et al., 2015). Weitere Analysen des Markers rs2303751 und seiner Funktion im *ISL1*-Locus sind daher indiziert (Zhang et al., 2017a).

7.1.2 Kandidatenanalyse von *SLC20A1*

SLC20A1 kodiert für einen Natriumphosphat-Symporter, der Phosphat in den intrazellulären Bereich transportiert. Dort wird das Phosphat dann zur Aufrechterhaltung von zellulären Funktionen wie Metabolismus, Signaltransduktion und Nukleinsäure- und Lipidsynthese absorbiert. *SLC20A1* scheint ebenfalls essenziell für die embryonale Entwicklung zu sein. Das homologe Protein *SLC20A2* kann im Tiermodell den Verlust von *SLC20A1* nicht kompensieren

(Beck et al., 2010). Wallingford und Giachelli haben im Jahr 2014 berichtet, dass *Slc20a1*-KO-Mäuseembryonen bei E12.5 versterben, als Hinweis darauf, dass das Protein SLC20A1 eine wesentliche Rolle bei der embryonalen Entwicklung spielt. Der Zeitrahmen (E9.5 bis E12.5) in welchem *SLC20A1*-KO-Mäuse versterben entspricht dem embryonalen Zeitfenster, für das bereits zuvor eine abnormale Dottersackgefäßentwicklung mit einer hohen Letalität bei Mäuseembryonen assoziiert war (Wallingford and Giachelli, 2014). Gemäß den Beobachtungen von Wallingford und Giachelli (2014) ist SLC20A1 notwendig für die Phosphat-vermittelte Endozytose über Mikroautophagie im viszeralem Endoderm des Dottersackes. Wie *SLC20A1* eine Rolle bei der Entwicklung von Harnwegen und Harnblasen aus dem viszeralem Endoderm des Dottersackes spielt, war bis dato unklar.

In der vorliegenden Arbeit wurden in drei unabhängigen Patienten insgesamt drei bisher nicht beschriebene Varianten in *SLC20A1* in 526 BEEK-Patienten (E: 35; KBE: 426; KE: 65) identifiziert. Die beiden neuen Varianten p.G237R und p.V298A wurden in den Patienten als *de novo* bestätigt. Die dritte Variante p.K441Q war mütterlich vererbt. Die Mutter wiederum hatte eine leichte Form des BEEK (übergroße Symphysenlücke). Anhand von UniProt wurde der Grad der Konservierung der identifizierten Missense-Mutationen im Protein SLC20A1 überprüft (Tabelle 26). Dabei zeigt sich, dass die Position der *de novo* Mutation p.G237R in anderen Vertebraten hochkonserviert ist. Nach dbNSFP wurde diese Missense-Mutation von acht Prädiktionsalgorithmen als krankheitsverursachend klassifiziert und nach „ACMG Standards and Guidelines“ wurde diese Mutation als pathogen vorgenommen. Die weiteren zwei Varianten p.V298A und p.K441Q sind ebenfalls hochkonserviert und wurden von einem bzw. vier Prädiktionsalgorithmen als krankheitsverursachend klassifiziert. Nach „ACMG Standards and Guidelines“ wurden die zwei Varianten p.V298A als pathogen und p.K441Q als vermutlich pathogen klassifiziert.

Tabelle 26 Darstellung des Grads der Konservierung der identifizierten Varianten mithilfe von UniProt

	p.G237R	p.V298A	p.K441Q
Human (Q8WUM9)	L I S V G C A V	N K H P V S E V	E G E Q K G E E
MOUSE (Q61609)	- - - - - - - -	H R N - - - - -	- - - - - - D -
RAT (Q9JJP0)	- - - - - - - -	- R N - - - - -	- - - - - - D -
FELCA (O97596)	- - - - - - - -	T R S - - - - -	- - - - - - - -
BOVIN (F1MLZ2)	- - - - - - - -	T R N - - - - -	- - - - - - - -
XENTER (Q5BL44)	- - - - - - - -	V K V - I L D M	D T - A R P D -
DANRE (Q6PFM1)	- - - I - F G I	x x x x x x x x	- S - F R A S -

-: gleiche Aminosäure an der Stelle; x: keine Aminosäure an der Stelle.

7.2 VATER/VACTERL-Assoziation/- ähnliche Phänotypen

Obwohl nur wenige große Studien veröffentlicht wurden, gibt es zahlreiche Hinweise, dass (*de novo*) CNVs mit der Expression von strukturellen Geburtsfehlern assoziiert sind (Brosens et al., 2016; Hilger et al., 2013; Southard et al., 2012). Je nach Geburtsfehler finden sich ursächliche *de novo* CNVs in bis zu 20% der Fälle (Brosens et al., 2016; Dworschak et al., 2015; Goldmuntz et al., 2011).

In der vorliegenden Arbeit wurde ein *de novo* Mikrodeletion mit einer Größe von 1,39Mb in der chromosomalen Region 17q12 (Chr17: 34.815.551-36.208.392) identifiziert. Diese Region entspricht der kritischen Region für das 17q12-Deletionssyndrom (Mitchel et al., 2016). Diese Deletion wurde in einer Patientin mit einem VATER/VACTERL-ähnlichen Phänotyp (A, R) identifiziert. Normalerweise umfasst das 17q12-Deletionssyndrom 15 Gene, eines davon ist das Gen *HNF1B*, welches den Hepatozyten-Kernfaktor 1 β kodiert. Zusätzlich zu Nierenanomalien sind die Mutationen und Deletionen in *HNF1B* mit genitalen Fehlbildungen, Autismus, Epilepsie, Gicht, Hypomagnesiämie, primärem Hyperparathyreoidismus, Leber- und Darmanomalien einschließlich Darmatresien assoziiert (Bockenbauer and Jaureguiberry, 2016). In DECIPHER wurde berichtet, dass ein Patient 284753 mit einer VATER/VACTERL-Assoziation (V, R, L) eine *de novo* Mikrodeletion in 17q12 hat, die komplett mit unser Mikrodeletion in 17q12 überlappt (Abb. 20) (Firth et al., 2009). Des Weiteren finden sich Berichte in der Literatur über *de novo* 17q12 Mikrodeletionen in Patienten mit VATER/VACTERL-ähnlichen Phänotypen (Faguer et al., 2011; Smigiel et al., 2014). Daraus schließen wir, dass die chromosomale Region 17q12 mit der Entstehung von Fehlbildungen aus dem VATER/VACTERL-Spektrum assoziiert ist (Zhang et al., 2017b).

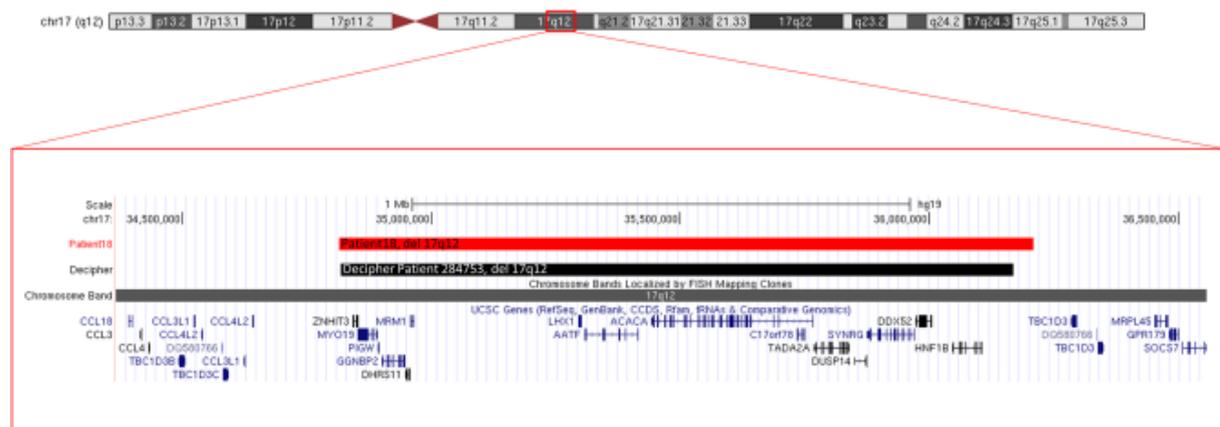


Abbildung 20 Darstellung der Mikrodeletion bei 17q12 in UCSC Genome Browser (Zhang et al., 2017b).

Für drei bestätigte Mikrodeletionen bei 5q23.1, 16q23.3, 22q11.21 und für eine Mikroduplikation bei 10q11.21 konnten ein *de novo* Ereignis nicht bestätigt werden, da nur ein Elternteil verfügbar war.

Nach bestem Wissen und Gewissen sind die Mikrodeletionen der chromosomalen Region 5q23.1 nicht mit VATER/VACTERL-ähnlichen Phänotypen assoziiert. Das einzige RefSeq-Gen, das in der hier gefundenen Mikrodeletion liegt ist *FAM170A*. Laut Mouse Genome Informatics (MGI) (<http://www.informatics.jax.org/>) ist *Fam170a* nur in murinen embryonalen Hoden, Spermatozyten und Sertoli-Zellen exprimiert. Da keines dieser Gewebe zu den klassischen VATER/VACTERL-Organsystemen gehört, scheint seine Beteiligung an der Entstehung von VATER/VACTERL-Phänotypen im Menschen unwahrscheinlich. Allerdings wurde vor kurzem eine überlappende Deletion in der chromosomalen Region 5q23.1 bei einem Patienten mit zerebellärer Hypoplasie identifiziert (Sajan et al., 2013). Daher kann nicht ausgeschlossen werden, dass die hier beobachtete Deletion in den chromosomalen Region 5q23.1 zur Expression des VATER/VACTERL-ähnlichen Phänotypen bei Patienten 5 beigetragen hat (Zhang et al., 2017b).

CDH13 ist das einzige RefSeq-Gen, das in der hier identifizierten Mikrodeletion 16q23.3 kartiert. Diese Deletion betrifft jedoch nur einen Teil des Intron-5 des Gens *CDH13*. Gleichwohl wird *Cdh13* nach MGI in fast allen VATER/VACTERL-relevanten Organsystemen während der Mausembryogenese exprimiert, was darauf hindeutet, dass heterozygote Deletionen des Gens *CDH13* im Menschen mit der Expression von VATER/VACTERL-ähnlichen Phänotypen assoziiert sein könnte (Zhang et al., 2017b). In der Literatur findet sich ein Fallbericht über einen Patienten mit globaler Entwicklungsverzögerung, Trigonozephalie, Portokaval-Shunt und Nephrokalzinose, der eine Deletion 16q23.1-q23.3 hat (Jobling et al., 2013). Dies deutet darauf hin, dass die Region bzw. das Gen *CDH13* eine Rolle in der Embryogenese der Menschen spielt.

Im Gegensatz zu den Deletionen in 5q23.1 und 16q23.3 wurden die *de novo* oder vererbten Deletionen in 22q11.21 zuvor mit VATER/VACTERL-ähnlichen Phänotypen assoziiert (Digilio et al., 1999; McDonald-McGinn et al., 2015). Die hier berichtete Patientin 20 mit 22q11.21-Mikrodeletion hat einen VATER/VACTERL-ähnlichen Phänotypen mit anorektaler Fehlbildung und Nierenanomalien. Sowohl anorektale Fehlbildungen als auch Nierenanomalien wurden

eine Ultraschalluntersuchung der Nieren. Daher können wir nicht ausschließen, dass irgendein Elternteil eine milde Ausprägung eines VATER/VACTERL-ähnlichen Phänotypen hat. In diesem Fall könnte einer der seltenen vererbten CNVs doch ein mit der Ausprägung der VATER/VACTERL-Assoziation assoziierter CNV sein (Zhang et al., 2017b).

In diesem Zusammenhang konnten in der aktuellen Arbeit zwei seltene aber vererbte Mikrodeletionen identifiziert werden, die in der chromosomalen Region 13q kartieren und vollständig mit der von Dworschak et al. (2013) beschriebenen *de novo* Deletion 13q31.1-q34, bei einem Patienten mit VATER/VACTERL-Assoziation (V, A, R, L) überlappen (Abb. 22). Außerdem hat die Patientin 7 mit anorektaler Fehlbildung und Speiseröhrenanomalien eine Deletion in der chromosomalen Region 7q35, die von ihrem „gesunden“ Vater vererbt wurde. Diese Deletion zeigt eine vollständige Überlappung mit einer als krankheitsverursachenden distalen *de novo* 7q-Deletion bei einem Patienten mit „V, A, TE“ von Brosens et al. (2016) (Abb. 23). Allerdings bleibt es ungewiss, ob diese Mikrodeletionen den gleichen Effekt haben, wie die großen von Dworschak et al. (2013) und Brosens et al. (2016) beschriebenen Deletionen mit einer Größe von 28,5Mb bzw. 15Mb (Zhang et al., 2017b).

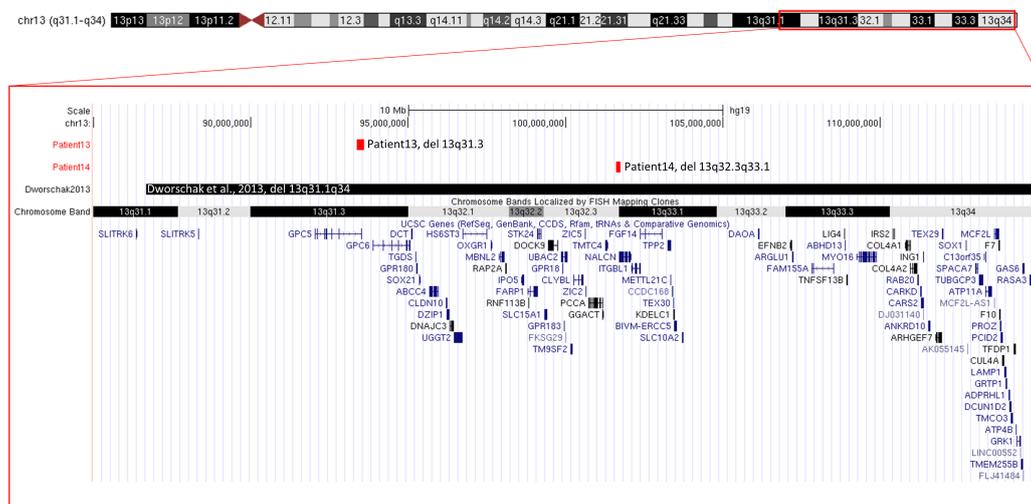


Abbildung 22 Darstellung der Überlappungen von 13q31.3del bei Patienten 13 und 13q32.3-q33.1del bei Patienten 14 mit der Deletion in 13q31.1-q34 von Dworschak et al. (2013) (Zhang et al., 2017b).

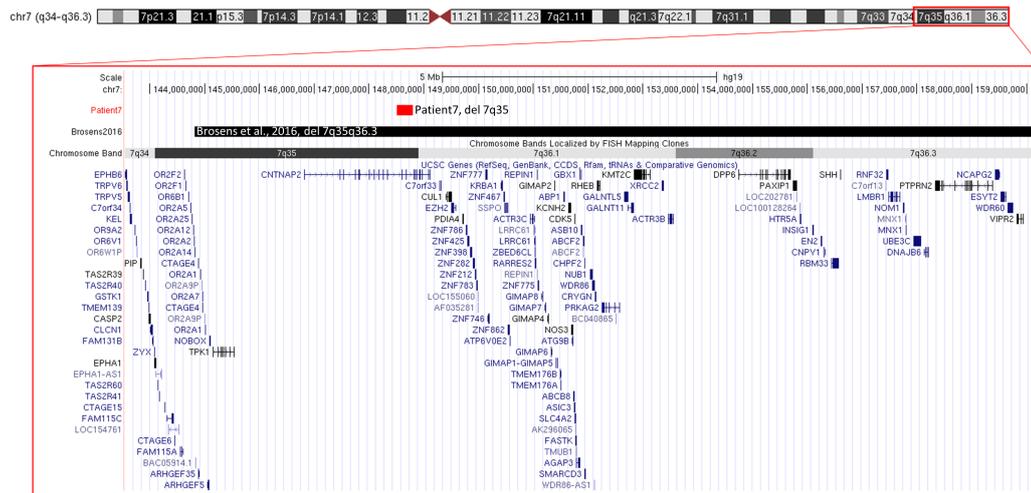


Abbildung 23 Darstellung der Überlappungen von 7q35del bei Patientin 7 mit der Deletion in 7q35-q36.3 von Brosens et al. (2016) (Zhang et al., 2017b).

7.3 Ösophagusatresien mit und ohne tracheo-ösophageale Fistel (EA/TEF)

7.3.1 Whole-Exome Sequencing (WES) bei Patienten mit Ösophagusatresien mit und ohne tracheo-ösophagealer Fistel (EA/TEF)

In dieser Arbeit durchgeführte WES Analyse ergab zunächst 26 vermeintliche *de novo* Mutationen in 31 Eltern-Kind-Trios. 24 dieser Mutationen konnten mittels Sanger Sequenzierung bestätigt werden. Davon wurden 14 bisher nicht beschrieben (nicht vorhanden in: dbSNP, ExAC Browser und 1000 Genomes (<http://phase3browser.1000genomes.org/index.html>)). Drei der 14 Varianten liegen in Genen, die zuvor mit angeborenen Anomalien der Speiseröhre assoziiert waren, z.B. *CHD7* (CHARGE-Syndrom), *FANCB* (Fanconi-Anämie) und *TRPS1* (Tricho-Rhino-Phalangealen Syndrome).

***CHD7* (chromodomain helicase DNA binding protein 7)**

Das CHARGE-Syndrom ist komplexe Multisystemerkrankung mit angeborenen Fehlbildungen und Entwicklungsstörungen in verschiedenen Organsystemen. Dieses Syndrom manifestiert sich äußerst variabel und wird durch Mutationen in *CHD7* verursacht (Corsten-Janssen et al., 2016). Ungefähr 19% der Patienten mit CHARGE-Syndrom und Mutationen in *CHD7* haben eine EA/TEF (Zentner et al., 2010). Mittels dbNSFP wurde die in dieser Arbeit identifizierte

Variante p.A1396G im Gen *CHD7* von acht In-Silico-Prädiktionsalgorithmen als pathologisch klassifiziert. Aufgrund des *de novo* Statuses und der Beurteilung durch die In-Silico-Prädiktionsalgorithmen gehen wir davon aus, dass diese Variante in *CHD7* zur Ausbildung der EA/TEF in unserem Patienten 167_501 führte.

***FANCB* (Fanconi anemia complementation group B)**

EA/TEF und die anderen gastrointestinalen Atresien sind eine gut dokumentierte Komponente von Fanconie-Anämie (FA). Gastrointestinale Atresien wurden in bis zu 14% der Patienten mit FA berichtet (Dokal, 2000). Nichts desto trotz ist der Anteil der Patienten mit einer FA unter allen Patienten mit einer EA/TEF gering. In der vorliegenden Arbeit wurde eine *de novo* Variante p.R261Q in *FANCB* in einer Patientin 46_501 identifiziert. Diese Variante wurde von neuen In-Silico-Prädiktionsalgorithmen von dbNSFP als benigne charakterisiert. Außerdem zeigt diese Patientin eine normale X-Chromosom-Inaktivierung. Demzufolge wurde diese Variante in *FANCB* als nicht pathologisch in der Patientin 46_501 gekennzeichnet.

***TRPS1* (transcriptional repressor GATA binding 1)**

Mutationen in *TRPS1* sind mit dem Auftreten der Tricho-Rhino-Phalangealen Syndrome (TRPS) assoziiert. TRPS-Patienten mit Mutationen in *TRPS1* haben gehäuft angeborene Herzfehler und gastroösophagealen Reflux (Maas et al., 2015). In unserem Patienten 288_501 zeigte sich eine bisher nicht beschriebene nonsense Mutation p.R544* in *TRPS1*. Die gleiche Mutation wurde zuvor bei drei TRPS-Patienten beobachtet. Anders als bei unserem Patienten 288_501 zeigten diese zuvor beschriebenen Patienten keine Herzfehler bzw. Fehlbildungen der Speise- oder Luftröhre (nach persönlichem Kontakt mit Herrn Raoul C. Hennekam). Interessanterweise zeigt *Trps1* einen durchweg hohen Expressionswert von 7,24 und logFC von 3,34 an den Tagen E8.5 und 12.5 in Mäuseembryonen im Bereich des Ösophagus. Zwischen beiden Tagen ist *Trps1* differentiell exprimiert mit einer deutlich stärkeren Expression an E12.5, so dass davon auszugehen ist, dass die hier vorliegende Mutation in *TRPS1* ursächlich an der Expression einer EA/TEF in unserem Patienten beteiligt ist.

Insgesamt zeigten drei von 23 Kandidatengen keine Expression während der Entwicklung des Ösophagus. Damit deckt sich unsere Beobachtung mit dem MGI-Datensatz (<http://www.Informatics.jax.org>). Interessanterweise gibt es 19 unserer 23 Kandidatengene keine Expressionsdaten in MGI, so dass die hier beschriebenen Expressionsmuster neu sind.

7.3.2 Kandidatengenanalyse von *EEF1D*

Da *Eef1d* die höchste durchschnittliche Expression während zu den untersuchten Zeitpunkten E8.5 und 12.5 im Bereich des sich entwickelnden Ösophagus in Mäuseembryonen zeigte und wir zudem eine trunkierende Neumutation in *EEF1D* in unserem Patienten identifizieren konnten, priorisierten wir *EEF1D* als Kandidatengen zur Re-Sequenzierung in unserer EA/TEF-Kohorte. Bislang wurden jedoch nur wenige Studien zu diesem Gen bzw. seinem Proteinprodukt publiziert. Elongationsfaktor 1 delta (*EEF1D*) ist ein Teil vom eukaryotischen Elongationsproteinkomplex (EF-1), welches den enzymatischen Transport von Aminoacyl-tRNAs an das Ribosom vermittelt (Riis et al., 1990). Ogawa et al. (2004) konnten zeigen, dass eine aberrante Expression von EF-1 das Zellwachstum stark beeinflusst (Ogawa et al., 2004). Ogawa et al. (2004) hat die Expression von *EEF1D* in malignen und benignen ösophagealem Gewebe von 52 Patienten mit Ösophaguskarzinom quantifiziert. Die Expression von EF-1 delta mRNA war in malignem Gewebe signifikant höher als im benignen Randgewebe (Ogawa et al., 2004). Es könnte sein, dass *EEF1D* eine Rolle bei Proliferation oder Apoptose spielt, da diese Prozesse bei der Entstehung von (Ösophagus-)Karzinomen eine große Rolle spielen. Bei unserem Patienten 4_501 wurde eine Stopmutation p.R292* in *EEF1D* identifiziert. Durch diese Stopmutation könnte die Translation der EF-1 delta mRNA früher abgebrochen werden, dadurch könnte die Genexpression reduziert sein. Eine kausale Rolle dieser Variante in der Entwicklung einer EA/TEF im Patienten 4_501 erscheint wahrscheinlich. Da wir in unserer Kohorte von 192 Patienten mit EA/TEF keine weitere kodierende Variante in *EEF1D* gefunden haben, ist davon auszugehen, dass Varianten in *EEF1D* eine seltene Ursache von EA/TEF ist. Welche Rolle *EEF1D* genau in der Entwicklung der EA/TEF spielt, konnte unsere Studie nicht beantworten und muss in zukünftigen experimentellen bzw. funktionellen Studien näher untersucht werden.

8 Zusammenfassung

Urorektale und gastrointestinale Malformationen sind menschliche Entwicklungsdefekte, die sich mit einem breiten phänotypischen Spektrum manifestieren und genetisch sehr heterogen sind. Die Ätiologie ist in den meisten Fällen bisher unbekannt. Das übergeordnete Ziel der hier vorgelegten Arbeit ist es, die genetischen Ursachen urorektales und gastrointestinales Fehlbildungen zu identifizieren und aufzuklären.

In der vorliegenden Arbeit lag der Fokus zuerst auf dem Blasenektrophie-Epispadie-Komplex (BEEK). In einer vorangegangenen Studie wurden die *ISL1*-Region als Risikolokus und weitere 11 nominant signifikante Regionen für multifaktorielle KBE-Phänotypen identifiziert. In der hier vorliegenden Arbeit wurden diese Regionen im Rahmen einer Replikationsstudie untersucht. Für die Replikationsstudie lagen sieben verschiedene Kollektive zur Verfügung, einschließlich 92 nordamerikanischer Eltern-Kind-Trios und 268 KBE-Patienten aus Australiern (n=31), Briten (n=40), Deutschen (n=7), Italienern (n=39), Spaniern (n=35) und Schweden (n=116) mit europäischer Herkunft und 1.354 ethnisch angeglichene Kontrollen. Insgesamt wurden 12 Marker für die 12 Regionen ausgewählt. Die *ISL1*-Region wurde erfolgreich in allen sieben Kohorten repliziert. Die anderen 11 Regionen haben das genomweit signifikante Niveau trotz vergrößerter Effekte nicht erreicht.

Durch die vormals durchgeführte Exom-Analyse in acht KE-Eltern-Kind-Trios wurde ein Risikogen *SLC20A1* identifiziert. Das Risikogen *SLC20A1* wurde in 526 BEEK-Patienten (E: 35; KBE: 426; KE: 65) re-sequenziert. In *SLC20A1* wurden insgesamt drei neue Varianten in drei unabhängigen Familien identifiziert. Von den drei Varianten wurden zwei als *de novo* bestätigt. Die dritte Variante wurde in einem Patienten und seiner leicht betroffenen Mutter gefunden.

Ein weiterer Fokus der vorliegenden Arbeit hatte die VATER/VACTERL-Assoziation im Blick. Die VATER/VACTERL-Assoziation beschreibt die seltene Assoziation verschiedener Fehlbildungen. Dabei wird die Diagnose klinisch gestellt, wenn drei der folgenden Fehlbildungen vorliegen: angeborene Defekte der Wirbelsäule (engl. vertebral, V), ARM (Analatresie, A), Fehlbildungen des Herzens (engl. cardiac, C), tracheo-ösophageale Fistel mit oder ohne Ösophagusatresie (engl. tracheoesophageal fistula with or without esophageal atresia, TE), Anomalien der Nieren (engl. renal, R) und Defekte der (radialen) Gliedmaßen (engl. limb, L). Wenn zwei der

Fehlbildungen vorliegen, kann man von einem VATER/VACTERL-ähnlichen Phänotypen sprechen.

In der vorliegenden Studie wurde eine Array-basierte Analyse zu ursächlichen Kopienzahlveränderungen (CNV) in 75 Patienten mit VATER/VACTERL-Assoziation und 40 Patienten mit VATER/VACTERL-ähnlichem Phänotyp durchgeführt. Insgesamt wurden 13 Mikrodeletionen und sieben Mikroduplikationen in 20 unabhängigen Patienten identifiziert, die nicht in gesunden Kontrollen zu finden waren. Dabei konnte eine Mikrodeletion in der chromosomalen Region 17q12 als *de novo* bestätigt werden. Für die Mikrodeletionen der chromosomalen Region 5q23.1, 16q23.3, und 22q11.21, und die Mikroduplikation der chromosomalen Region 10q11.21 konnte nicht überprüft werden, ob diese *de novo* vorliegen, da nur DNA eines Elternteils zur Verfügung stand. Das vorhandene Elternteil trug den CNV nicht. Da Mikrodeletionen der Region 22q11.21 bereits zuvor mit Fehlbildungen und Kombinationen von Fehlbildungen aus dem VATER/VACTERL-Spektrum assoziiert waren, gehen wir davon aus, dass diese Mikrodeletion hier ursächlich ist. Die restlichen 15 CNVs wurden jeweils von einem gesunden Elternteil vererbt. Die vorliegenden Ergebnisse deuten darauf hin, dass (seltene) CNVs zur Entstehung der VATER/VACTERL-Assoziation oder VATER/VACTERL-ähnlichen Phänotypen beitragen können. Weitere Studien sind erforderlich, um zu klären, wie der Verlust bzw. Zugewinn der jeweiligen CNVs dazu beiträgt, dass die CNV-tragenden Patienten Fehlbildungen aus dem VATER/VACTERL-Spektrum ausbilden.

Im letzten Teil der vorliegenden Arbeit lag der Fokus auf angeborenen Ösophagusatresien mit und ohne tracheo-ösophagealer Fistel (EA/TEF). Um kleinere ursächliche Neumutationen im Bereich von Punktmutationen identifizieren zu können wurden mittels Whole-Exome Sequencing (WES) 31 Eltern-Kinder-Trios untersucht. Dabei konnten 24 Varianten in 23 Genen und 18 unabhängigen Patienten mit Hilfe von Sanger Sequenzierung bestätigt werden. Um die Kandidatengene weiter zu priorisieren, wurden Transkriptom-Analysen im sich entwickelnden Ösophagus bei Mäuseembryonen an den Tagen E8.5 (TS13) und E12.5 (TS21) durchgeführt. Des Weiteren wurden die Varianten mittels Vorhersagealgorithmen klassifiziert. In Zusammenschau aller Informationen wurde die trunkierende Variante in *EEF1D* priorisiert und *EEF1D* als Kandidatengenen in unserem EA/TEF Kollektiv re-sequenziert. Die Sangersequenzierung von 192 weiteren EA/TEF-Patienten konnte keine weitere Variante identifizieren. In einem nächsten Schritt soll *EEF1D* und weitere priorisierte Kandidatengene

in der gesamten EA/TEF-Kohorte mittels smMIPs (engl. single molecule molecular inversion probes) re-sequenziert werden, um weitere Mutationsträger für die priorisierten Gene zu identifizieren. Die smMIP-basierte Re-sequenzierung erlaubt es viele Gene und Patienten mit geringerem Kostenaufwand zu untersuchen.

Zusammenfassend hat die vorgelegte Arbeit dazu beigetragen bisher nicht bekannte genetische Faktoren für die untersuchten Fehlbildungen zu identifizieren und zu bestätigen. Dadurch hat die vorgelegte Arbeit einen wichtigen Beitrag zum besseren Verständnis der Ursachen urorektaler und gastrointestinaler Fehlbildungen und letztendlich der Entwicklung dieser Organstrukturen geleistet. Weitere Studien sind für die endgültige Aufklärung der genetischen Ursachen von BEEK, VATER/VACTERL-Assoziation und EA/TEF erforderlich. Zukünftige GWAS an größeren Kohorten werden weitere genetische Regionen mit Suzeptibilität für multifaktorielle Phänotypen identifizieren und ggf. Hinweise auf Gen-Umweltfaktor-Interaktionen liefern.

Darüber hinaus werden umfassende Analysen mittels WES, Whole-Genome Sequencing und smMIPs Varianten zur Identifizierung von weiteren hochpenetranten Varianten führen, die im Einzelfall einen direkten Rückschluss auf die Funktion des betreffenden Genes in der Expression der entsprechenden Fehlbildung zulassen. Die so identifizierten Regionen und Kandidatengene werden dann in Vertebratenmodellen (z.B. Zebrafisch) in ihrer Funktion morphologisch und funktionell detailliert charakterisiert.

Referenzen

- Alamo, L., Meyrat, B.J., Meuwly, J.-Y., Meuli, R.A., and Gudinchet, F. (2013). Anorectal Malformations: Finding the Pathway out of the Labyrinth. *RadioGraphics* 33, 491–512.
- AMPure XP Protocol_000387v001 AMPure XP - System zu Aufreinigung von PCR-Produkten.
- Antonarakis, S.E. (2010). Human Genome Sequence and Variation. In Vogel and Motulsky's Human Genetics, M.R. Speicher, A.G. Motulsky, and S.E. Antonarakis, eds. (Springer Berlin Heidelberg), pp. 31–53.
- Arrington, C.B., Patel, A., Bacino, C.A., and Bowles, N.E. (2010). Haploinsufficiency of the LIM domain containing preferred translocation partner in lipoma (LPP) gene in patients with tetralogy of Fallot and VACTERL association. *Am. J. Med. Genet. A.* 152A, 2919–2923.
- Attiyeh, E.F., Diskin, S.J., Attiyeh, M.A., Mossé, Y.P., Hou, C., Jackson, E.M., Kim, C., Glessner, J., Hakonarson, H., Biegel, J.A., et al. (2009). Genomic copy number determination in cancer cells from single nucleotide polymorphism microarrays based on quantitative genotyping corrected for aneuploidy. *Genome Res.* 19, 276–283.
- Bamshad, M.J., Ng, S.B., Bigham, A.W., Tabor, H.K., Emond, M.J., Nickerson, D.A., and Shendure, J. (2011). Exome sequencing as a tool for Mendelian disease gene discovery. *Nat. Rev. Genet.* 12, 745–755.
- Barbazuk, W.B., Korf, I., Kadavi, C., Heyen, J., Tate, S., Wun, E., Bedell, J.A., McPherson, J.D., and Johnson, S.L. (2000). The Syntenic Relationship of the Zebrafish and Human Genomes. *Genome Res.* 10, 1351–1358.
- Beaudoin, S., Barbet, P., and Bargy, F. (2004). Pelvic development in the rabbit embryo: implications in the organogenesis of bladder exstrophy. *Anat. Embryol. (Berl.)* 208, 425–430.
- Beck, L., Leroy, C., Beck-Cormier, S., Forand, A., Salaün, C., Paris, N., Bernier, A., Ureña-Torres, P., Prié, D., Ollero, M., et al. (2010). The Phosphate Transporter PiT1 (Slc20a1) Revealed As a New Essential Gene for Mouse Liver Development. *PLOS ONE* 5, e9148.
- Beetz, R., Mannhardt-Laakmann, W., and Schofer, O. (1998). *Kinderurologische Sprechstunde*. Wissenschaftliche Verlagsgesellschaft, Stuttgart.
- Belloni, E., Martucciello, G., Verderio, D., Ponti, E., Seri, M., Jasonni, V., Torre, M., Ferrari, M., Tsui, L.-C., and Scherer, S.W. (2000). Involvement of the HLXB9 Homeobox Gene in Currarino Syndrome. *Am. J. Hum. Genet.* 66, 312–319.
- Bertier, G., Héту, M., and Joly, Y. (2016). Unsolved challenges of clinical whole-exome sequencing: a systematic literature review of end-users' views. *BMC Med. Genomics* 9.
- Bhatnagar, S. (2015). Anorectal Malformations (Part 1). *J. Neonatal Surg.* 4.
- Biemar, F., Argenton, F., Schmidtke, R., Epperlein, S., Peers, B., and Driever, W. (2001). Pancreas development in zebrafish: early dispersed appearance of endocrine hormone expressing cells and their convergence to form the definitive islet. *Dev. Biol.* 230, 189–203.
- Bill, B.R., Petzold, A.M., Clark, K.J., Schimmenti, L.A., and Ekker, S.C. (2009). A Primer for Morpholino Use in Zebrafish. *Zebrafish* 6, 69–77.

- Bischoff, A., Levitt, M.A., and Peña, A. (2013). Update on the management of anorectal malformations. *Pediatr. Surg. Int.* *29*, 899–904.
- Blake, J.A., Eppig, J.T., Kadin, J.A., Richardson, J.E., Smith, C.L., Bult, C.J., and the Mouse Genome Database Group (2017). Mouse Genome Database (MGD)-2017: community knowledge resource for the laboratory mouse. *Nucleic Acids Res.* *45*, D723–D729.
- Bockenbauer, D., and Jaureguiberry, G. (2016). HNF1B-associated clinical phenotypes: the kidney and beyond. *Pediatr. Nephrol.* *31*, 707–714.
- Bohring, A. (2002). OEIS complex, VATER, and the ongoing difficulties in terminology and delineation. *Am. J. Med. Genet.* *107*, 72–76.
- van Bokhoven, H., Celli, J., van Reeuwijk, J., Rinne, T., Glaudemans, B., van Beusekom, E., Rieu, P., Newbury-Ecob, R.A., Chiang, C., and Brunner, H.G. (2005). MYCN haploinsufficiency is associated with reduced brain size and intestinal atresias in Feingold syndrome. *Nat. Genet.* *37*, 465–467.
- Bonnot, O., Vollset, S.E., Godet, P.F., D'Amato, T., and Robert, E. (2001). Maternal exposure to lorazepam and anal atresia in newborns: results from a hypothesis-generating study of benzodiazepines and malformations. *J. Clin. Psychopharmacol.* *21*, 456–458.
- Boocock, G.R., and Donnai, D. (1987). Anorectal malformation: familial aspects and associated anomalies. *Arch. Dis. Child.* *62*, 576–579.
- Botstein, D., and Risch, N. (2003). Discovering genotypes underlying human phenotypes: past successes for mendelian disease, future approaches for complex disease. *Nat. Genet.* *33 Suppl*, 228–237.
- Boyadjiev, S.A., Dodson, J.L., Radford, C.L., Ashrafi, G.H., Beaty, T.H., Mathews, R.I., Broman, K.W., and Gearhart, J.P. (2004). Clinical and molecular characterization of the bladder exstrophy-epispadias complex: analysis of 232 families. *BJU Int.* *94*, 1337–1343.
- Brosens, E., Eussen, H., van Bever, Y., van der Helm, R.M., Ijsselstijn, H., Zaveri, H.P., Wijnen, R., Scott, D.A., Tibboel, D., and de Klein, A. (2013). VACTERL Association Etiology: The Impact of de novo and Rare Copy Number Variations. *Mol. Syndromol.* *4*, 20–26.
- Brosens, E., Ploeg, M., van Bever, Y., Koopmans, A.E., Ijsselstijn, H., Rottier, R.J., Wijnen, R., Tibboel, D., and de Klein, A. (2014a). Clinical and etiological heterogeneity in patients with tracheo-esophageal malformations and associated anomalies. *Eur. J. Med. Genet.* *57*, 440–452.
- Brosens, E., de Jong, E.M., Barakat, T.S., Eussen, B.H., D'haene, B., De Baere, E., Verdin, H., Poddighe, P.J., Galjaard, R.-J., Gribnau, J., et al. (2014b). Structural and numerical changes of chromosome X in patients with esophageal atresia. *Eur. J. Hum. Genet. EJHG* *22*, 1077–1084.
- Brosens, E., Marsch, F., de Jong, E.M., Zaveri, H.P., Hilger, A.C., Choinitzki, V.G., Hölscher, A., Hoffmann, P., Herms, S., Boemers, T.M., et al. (2016). Copy number variations in 375 patients with oesophageal atresia and/or tracheoesophageal fistula. *Eur. J. Hum. Genet. EJHG* *24*, 1715–1723.
- Brown, A.K., Roddam, A.W., Spitz, L., and Ward, S.J. (1999). Oesophageal atresia, related malformations, and medical problems: A family study. *Am. J. Med. Genet.* *85*, 31–37.
- Brunner, H.G., and van Bokhoven, H. (2005). Genetic players in esophageal atresia and tracheoesophageal fistula. *Curr. Opin. Genet. Dev.* *15*, 341–347.

- Burton, P.R., Clayton, D.G., Cardon, L.R., Craddock, N., Deloukas, P., Duncanson, A., Kwiatkowski, D.P., McCarthy, M.I., Ouwehand, W.H., Samani, N.J., et al. (2007). Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* 447, 661–678.
- Busa, T., Panait, N., Chaumoitre, K., Philip, N., and Missirian, C. (2016). Esophageal atresia with tracheoesophageal fistula in a patient with 7q35-36.3 deletion including SHH gene. *Eur. J. Med. Genet.* 59, 546–548.
- Bush, W.S., and Moore, J.H. (2012). Chapter 11: Genome-Wide Association Studies. *PLOS Comput. Biol.* 8, e1002822.
- Cai, C.-L., Liang, X., Shi, Y., Chu, P.-H., Pfaff, S.L., Chen, J., and Evans, S. (2003). Isl1 identifies a cardiac progenitor population that proliferates prior to differentiation and contributes a majority of cells to the heart. *Dev. Cell* 5, 877–889.
- Carey, J.C., Greenbaum, B., and Hall, B.D. (1978). The OEIS complex (omphalocele, exstrophy, imperforate anus, spinal defects). *Birth Defects Orig. Artic. Ser XIV*: 253–268.
- Caton, A.R., Bloom, A., Druschel, C.M., and Kirby, R.S. (2007). Epidemiology of bladder and cloacal exstrophies in New York State, 1983-1999. *Birt. Defects Res. A. Clin. Mol. Teratol.* 79, 781–787.
- Chanock, S.J., Manolio, T., Boehnke, M., Boerwinkle, E., Hunter, D.J., Thomas, G., Hirschhorn, J.N., Abecasis, G., Altshuler, D., Bailey-Wilson, J.E., et al. (2007). Replicating genotype–phenotype associations. *Nature* 447, 655–660.
- Cheng, W., Jacobs, W.B., Zhang, J.J.R., Moro, A., Park, J.-H., Kushida, M., Qiu, W., Mills, A.A., and Kim, P.C.W. (2006). DeltaNp63 plays an anti-apoptotic role in ventral bladder development. *Dev. Camb. Engl.* 133, 4783–4792.
- Cho, S., Moore, S.P., and Fangman, T. (2001). One hundred three consecutive patients with anorectal malformations and their associated anomalies. *Arch. Pediatr. Adolesc. Med.* 155, 587–591.
- Chung, B., Shaffer, L.G., Keating, S., Johnson, J., Casey, B., and Chitayat, D. (2011). From VACTERL-H to heterotaxy: variable expressivity of ZIC3-related disorders. *Am. J. Med. Genet. A.* 155A, 1123–1128.
- Cinti, R., Priolo, M., Lerone, M., Gimelli, G., Seri, M., Silengo, M., and Ravazzolo, R. (2001). Molecular characterisation of a supernumerary ring chromosome in a patient with VATER association. *J. Med. Genet.* 38, E6.
- CleanSEQ Protocol_000600v032 Agencourt CleanSEQ - System zu Aufreinigung von Sequenzierprodukten (Dye Removal).
- Colella, S., Yau, C., Taylor, J.M., Mirza, G., Butler, H., Clouston, P., Bassett, A.S., Seller, A., Holmes, C.C., and Ragoussis, J. (2007). QuantiSNP: an Objective Bayes Hidden-Markov Model to detect and accurately map copy number variation using SNP genotyping data. *Nucleic Acids Res.* 35, 2013–2025.
- Conrad, D.F., Pinto, D., Redon, R., Feuk, L., Gokcumen, O., Zhang, Y., Aerts, J., Andrews, T.D., Barnes, C., Campbell, P., et al. (2010). Origins and functional impact of copy number variation in the human genome. *Nature* 464, 704–712.
- Corsten-Janssen, N., van Ravenswaaij-Arts, C.M.A., and Kapusta, L. (2016). Congenital arch vessel anomalies in CHARGE syndrome: A frequent feature with risk for co-morbidity. *Int. J. Cardiol. Heart Vasc.* 12, 21–25.

- Crétolle, C., Pelet, A., Sanlaville, D., Zérah, M., Amiel, J., Jaubert, F., Révillon, Y., Baala, L., Munnich, A., Nihoul-Fékété, C., et al. (2008). Spectrum of HLXB9 gene mutations in Currarino syndrome and genotype-phenotype correlation. *Hum. Mutat.* *29*, 903–910.
- Cuschieri, A. (2002). Anorectal anomalies associated with or as part of other anomalies. *Am. J. Med. Genet.* *110*, 122–130.
- David, T.J., and O’Callaghan, S.E. (1975). Oesophageal atresia in the South West of England. *J. Med. Genet.* *12*, 1–11.
- Davies, B.R., and Giménez-Scherer, J.A. (2011). Two major patterns of nongenetic malformations are found at autopsy. *Pediatr. Dev. Pathol. Off. J. Soc. Pediatr. Pathol. Paediatr. Pathol. Soc.* *14*, 206–213.
- Depaepe, A., Dolk, H., and Lechat, M.F. (1993). The epidemiology of tracheo-oesophageal fistula and oesophageal atresia in Europe. EUROCAT Working Group. *Arch. Dis. Child.* *68*, 743–748.
- Di Gianantonio, E., Schaefer, C., Mastroiacovo, P.P., Cournot, M.P., Benedicenti, F., Reuvers, M., Occupati, B., Robert, E., Bellemin, B., Addis, A., et al. (2001). Adverse effects of prenatal methimazole exposure. *Teratology* *64*, 262–266.
- Digilio, M.C., Marino, B., Bagolan, P., Giannotti, A., and Dallapiccola, B. (1999). Microdeletion 22q11 and oesophageal atresia. *J. Med. Genet.* *36*, 137–139.
- Dokal, I. (2000). The genetics of Fanconi’s anaemia. *Best Pract. Res. Clin. Haematol.* *13*, 407–425.
- Dooley, K., and Zon, L.I. (2000). Zebrafish: a model system for the study of human disease. *Curr. Opin. Genet. Dev.* *10*, 252–256.
- Draaken, M., Reutter, H., Schramm, C., Bartels, E., Boemers, T.M., Ebert, A.-K., Rösch, W., Schröder, A., Stein, R., Moebus, S., et al. (2010). Microduplications at 22q11.21 are associated with non-syndromic classic bladder exstrophy. *Eur. J. Med. Genet.* *53*, 55–60.
- Draaken, M., Mughal, S.S., Pennimpede, T., Wolter, S., Wittler, L., Ebert, A.-K., Rösch, W., Stein, R., Bartels, E., Schmidt, D., et al. (2013). Isolated bladder exstrophy associated with a de novo 0.9 Mb microduplication on chromosome 19p13.12. *Birt. Defects Res. A. Clin. Mol. Teratol.* *97*, 133–139.
- Draaken, M., Baudisch, F., Timmermann, B., Kuhl, H., Kerick, M., Proske, J., Wittler, L., Pennimpede, T., Ebert, A.-K., Rösch, W., et al. (2014). Classic bladder exstrophy: Frequent 22q11.21 duplications and definition of a 414 kb phenocritical region. *Birt. Defects Res. A. Clin. Mol. Teratol.* *100*, 512–517.
- Draaken, M., Knapp, M., Pennimpede, T., Schmidt, J.M., Ebert, A.-K., Rösch, W., Stein, R., Utsch, B., Hirsch, K., Boemers, T.M., et al. (2015). Genome-wide association study and meta-analysis identify ISL1 as genome-wide significant susceptibility gene for bladder exstrophy. *PLoS Genet.* *11*, e1005024.
- Drummond, I.A., Majumdar, A., Hentschel, H., Elger, M., Solnica-Krezel, L., Schier, A.F., Neuhauss, S.C., Stemple, D.L., Zwartkuis, F., Rangini, Z., et al. (1998). Early development of the zebrafish pronephros and analysis of mutations affecting pronephric function. *Dev. Camb. Engl.* *125*, 4655–4667.
- Durinck, S., Moreau, Y., Kasprzyk, A., Davis, S., De Moor, B., Brazma, A., and Huber, W. (2005). BioMart and Bioconductor: a powerful link between biological databases and microarray data analysis. *Bioinforma. Oxf. Engl.* *21*, 3439–3440.

- Durinck, S., Spellman, P.T., Birney, E., and Huber, W. (2009). Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nat. Protoc.* *4*, 1184–1191.
- Dworschak, G.C., Draaken, M., Marcelis, C., de Blaauw, I., Pfundt, R., van Rooij, I.A.L.M., Bartels, E., Hilger, A., Jenetzky, E., Schmiedeke, E., et al. (2013). De novo 13q deletions in two patients with mild anorectal malformations as part of VATER/VACTERL and VATER/VACTERL-like association and analysis of EFNB2 in patients with anorectal malformations. *Am. J. Med. Genet. A.* *161*, 3035–3041.
- Dworschak, G.C., Draaken, M., Hilger, A.C., Schramm, C., Bartels, E., Schmiedeke, E., Grasshoff-Derr, S., Märzheuser, S., Holland-Cunz, S., Lacher, M., et al. (2015). Genome-wide mapping of copy number variations in patients with both anorectal malformations and central nervous system abnormalities. *Birt. Defects Res. A. Clin. Mol. Teratol.* *103*.
- Ebert, A.-K., Reutter, H., Ludwig, M., and Rösch, W.H. (2009). The Exstrophy-epispadias complex. *Orphanet J. Rare Dis.* *4*, 23.
- El-Gohary, Y., Gittes, G.K., and Tovar, J.A. (2010). Congenital anomalies of the esophagus. *Semin. Pediatr. Surg.* *19*, 186–193.
- El-Hattab, A.W., Skorupski, J.C., Hsieh, M.H., Breman, A.M., Patel, A., Cheung, S.W., and Craigen, W.J. (2010). OEIS complex associated with chromosome 1p36 deletion: a case report and review. *Am. J. Med. Genet. A.* *152A*, 504–511.
- Ensembl (2017). Homo sapiens - Ensembl genome browser 89. GRCh38.p10. 2017. http://www.ensembl.org/Homo_sapiens/Info/Annotation. Accessed 28 Jun 2017.
- Erzurumluoglu, A.M., Rodriguez, S., Shihab, H.A., Baird, D., Richardson, T.G., Day, I.N.M., and Gaunt, T.R. (2015). Identifying Highly Penetrant Disease Causal Mutations Using Next Generation Sequencing: Guide to Whole Process.
- Faguer, S., Decramer, S., Chassaing, N., Bellanné-Chantelot, C., Calvas, P., Beaufile, S., Bessenay, L., Lengelé, J.-P., Dahan, K., Ronco, P., et al. (2011). Diagnosis, management, and prognosis of HNF1B nephropathy in adulthood. *Kidney Int.* *80*, 768–776.
- Falardeau, J., Chung, W.C.J., Beenken, A., Raivio, T., Plummer, L., Sidis, Y., Jacobson-Dickman, E.E., Eliseenkova, A.V., Ma, J., Dwyer, A., et al. (2008). Decreased FGF8 signaling causes deficiency of gonadotropin-releasing hormone in humans and mice. *J. Clin. Invest.* *118*, 2822–2831.
- Feldkamp, M.L., Botto, L.D., Amar, E., Bakker, M.K., Bermejo-Sánchez, E., Bianca, S., Canfield, M.A., Castilla, E.E., Clementi, M., Csaky-Szunyogh, M., et al. (2011). Cloacal exstrophy: an epidemiologic study from the International Clearinghouse for Birth Defects Surveillance and Research. *Am. J. Med. Genet. C Semin. Med. Genet.* *157C*, 333–343.
- Felix, J.F., Steegers-Theunissen, R.P.M., de Walle, H.E.K., de Klein, A., Torfs, C.P., and Tibboel, D. (2007). Esophageal atresia and tracheoesophageal fistula in children of women exposed to diethylstilbestrol in utero. *Am. J. Obstet. Gynecol.* *197*, 38.e1-5.
- Felix, J.F., van Dooren, M.F., Klaassens, M., Hop, W.C.J., Torfs, C.P., and Tibboel, D. (2008). Environmental factors in the etiology of esophageal atresia and congenital diaphragmatic hernia: results of a case-control study. *Birt. Defects Res. A. Clin. Mol. Teratol.* *82*, 98–105.
- Felix, J.F., de Jong, E.M., Torfs, C.P., de Klein, A., Rottier, R.J., and Tibboel, D. (2009). Genetic and environmental factors in the etiology of esophageal atresia and/or tracheoesophageal fistula: an overview of the current concepts. *Birt. Defects Res. A. Clin. Mol. Teratol.* *85*, 747–754.

- Ferretti, E., Li, B., Zewdu, R., Wells, V., Hebert, J.M., Karner, C., Anderson, M.J., Williams, T., Dixon, J., Dixon, M.J., et al. (2011). A conserved Pbx-Wnt-p63-Irf6 regulatory module controls face morphogenesis by promoting epithelial apoptosis. *Dev. Cell* *21*, 627–641.
- Firth, H.V., Richards, S.M., Bevan, A.P., Clayton, S., Corpas, M., Rajan, D., Vooren, S.V., Moreau, Y., Pettett, R.M., and Carter, N.P. (2009). DECIPHER: Database of Chromosomal Imbalance and Phenotype in Humans Using Ensembl Resources. *Am. J. Hum. Genet.* *84*, 524–533.
- Fu, W., O'Connor, T.D., Jun, G., Kang, H.M., Abecasis, G., Leal, S.M., Gabriel, S., Rieder, M.J., Altshuler, D., Shendure, J., et al. (2013). Analysis of 6,515 exomes reveals the recent origin of most human protein-coding variants. *Nature* *493*, 216–220.
- Gambhir, L., Höller, T., Müller, M., Schott, G., Vogt, H., Detlefsen, B., Ebert, A.-K., Fisch, M., Beaudoin, S., Stein, R., et al. (2008). Epidemiological survey of 214 families with bladder exstrophy-epispadias complex. *J. Urol.* *179*, 1539–1543.
- Garcia-Barceló, M.-M., Wong, K.K., Lui, V.C., Yuan, Z., So, M., Ngan, E.S., Miao, X., Chung, P.H., Khong, P., and Tam, P.K. (2008). Identification of a HOXD13 mutation in a VACTERL patient. *Am. J. Med. Genet. A.* *146A*, 3181–3185.
- Gearhart, J. (1998). New potential for human embryonic stem cells. *Science* *282*, 1061–1062.
- GENCODE (2016). Human GENCODE version 26. 2016. <https://www.gencodegenes.org/stats/current.html>. Accessed 28 Jun 2017.
- Geneviève, D., de Pontual, L., Amiel, J., Sarnacki, S., and Lyonnet, S. (2007). An overview of isolated and syndromic oesophageal atresia. *Clin. Genet.* *71*, 392–399.
- Goldmuntz, E., Paluru, P., Glessner, J., Hakonarson, H., Biegel, J.A., White, P.S., Gai, X., and Shaikh, T.H. (2011). Microdeletions and Microduplications in Patients with Congenital Heart Disease and Multiple Congenital Anomalies. *Congenit. Heart Dis.* *6*, 592–602.
- Gunderson, K.L., Kruglyak, S., Graige, M.S., Garcia, F., Kermani, B.G., Zhao, C., Che, D., Dickinson, T., Wickham, E., Bierle, J., et al. (2004). Decoding Randomly Ordered DNA Arrays. *Genome Res.* *14*, 870–877.
- Guo, B., Sallis, R.E., Greenall, A., Petit, M.M.R., Jansen, E., Young, L., Van de Ven, W.J.M., and Sharrocks, A.D. (2006). The LIM domain protein LPP is a coactivator for the ETS domain transcription factor PEA3. *Mol. Cell. Biol.* *26*, 4529–4538.
- Guo, Y., He, J., Zhao, S., Wu, H., Zhong, X., Sheng, Q., Samuels, D.C., Shyr, Y., and Long, J. (2014). Illumina human exome genotyping array clustering and quality control. *Nat. Protoc.* *9*, 2643–2662.
- Haff, L.A., and Smirnov, I.P. (1997). Single-nucleotide polymorphism identification assays using a thermostable DNA polymerase and delayed extraction MALDI-TOF mass spectrometry. *Genome Res.* *7*, 378–388.
- Hagan, D.M., Ross, A.J., Strachan, T., Lynch, S.A., Ruiz-Perez, V., Wang, Y.M., Scambler, P., Custard, E., Reardon, W., Hassan, S., et al. (2000). Mutation analysis and embryonic expression of the HLXB9 Currarino syndrome gene. *Am. J. Hum. Genet.* *66*, 1504–1515.
- Hamid, C.H., Holland, A.J.A., and Martin, H.C.O. (2007). Long-term outcome of anorectal malformations: the patient perspective. *Pediatr. Surg. Int.* *23*, 97–102.

- Haraguchi, R., Mo, R., Hui, C., Motoyama, J., Makino, S., Shiroishi, T., Gaffield, W., and Yamada, G. (2001). Unique functions of Sonic hedgehog signaling during external genitalia development. *Dev. Camb. Engl.* *128*, 4241–4250.
- Haraguchi, R., Motoyama, J., Sasaki, H., Satoh, Y., Miyagawa, S., Nakagata, N., Moon, A., and Yamada, G. (2007). Molecular analysis of coordinated bladder and urogenital organ formation by Hedgehog signaling. *Dev. Camb. Engl.* *134*, 525–533.
- Hernández-García, A., Brosens, E., Zaveri, H.P., de Jong, E.M., Yu, Z., Namwanje, M., Mayle, A., Fernandes, C.J., Lee, B., Blazo, M., et al. (2012). Contribution of LPP copy number and sequence changes to esophageal atresia, tracheoesophageal fistula, and VACTERL association. *Am. J. Med. Genet. A.* *158A*, 1785–1787.
- Hilger, A., Schramm, C., Pennimpede, T., Wittler, L., Dworschak, G.C., Bartels, E., Engels, H., Zink, A.M., Degenhardt, F., Müller, A.M., et al. (2013). De novo microduplications at 1q41, 2q37.3, and 8q24.3 in patients with VATER/VACTERL association. *Eur. J. Hum. Genet.* *21*, 1377–1382.
- Hill, A.J., Teraoka, H., Heideman, W., and Peterson, R.E. (2005). Zebrafish as a Model Vertebrate for Investigating Chemical Toxicity. *Toxicol. Sci.* *86*, 6–19.
- Holden, S.T., Cox, J.J., Kesterton, I., Thomas, N.S., Carr, C., and Woods, C.G. (2006). Fanconi anaemia complementation group B presenting as X linked VACTERL with hydrocephalus syndrome. *J. Med. Genet.* *43*, 750–754.
- Hölscher, A.C., Laschat, M., Choinitzki, V., Zwink, N., Jenetzky, E., Münsterer, O., Kurz, R., Pauly, M., Brokmeier, U., Leutner, A., et al. (2017). Quality of Life after Surgical Treatment for Esophageal Atresia: Long-Term Outcome of 154 Patients. *Eur. J. Pediatr. Surg. Off. J. Austrian Assoc. Pediatr. Surg. Al Z. Kinderchir.* *27*, 443–448.
- Holschneider, A., Hutson, J., Peña, A., Beket, E., Chatterjee, S., Coran, A., Davies, M., Georgeson, K., Grosfeld, J., Gupta, D., et al. (2005). Preliminary report on the International Conference for the Development of Standards for the Treatment of Anorectal Malformations. *J. Pediatr. Surg.* *40*, 1521–1526.
- Howe, K., Clark, M.D., Torroja, C.F., Torrance, J., Berthelot, C., Muffato, M., Collins, J.E., Humphray, S., McLaren, K., Matthews, L., et al. (2013). The zebrafish reference genome sequence and its relationship to the human genome. *Nature* *496*, 498–503.
- ICBDMS (1987). Epidemiology of bladder exstrophy and epispadias: a communication from the International Clearinghouse for Birth Defects Monitoring Systems. *Teratology* *36*, 221–227.
- Ihrie, R.A., Marques, M.R., Nguyen, B.T., Horner, J.S., Papazoglu, C., Bronson, R.T., Mills, A.A., and Attardi, L.D. (2005). Perp is a p63-regulated gene essential for epithelial integrity. *Cell* *120*, 843–856.
- International Clearinghouse for Birth Defects Monitoring Systems (1987). Epidemiology of bladder exstrophy and epispadias: A communication from the international clearinghouse for birth defects monitoring systems. *Teratology* *36*, 221–227.
- International Human Genome Sequencing Consortium (2004). Finishing the euchromatic sequence of the human genome. *Nature* *431*, 931–945.
- Ioannides, A.S., and Copp, A.J. (2009). Embryology of oesophageal atresia. *Semin. Pediatr. Surg.* *18*, 2–11.

- Ives, E., Coffey, R., and Carter, C.O. (1980). A family study of bladder exstrophy. *J. Med. Genet.* *17*, 139–141.
- Jenetzky, E. (2007). Prevalence estimation of anorectal malformations using German diagnosis related groups system. *Pediatr. Surg. Int.* *23*, 1161–1165.
- Jobling, R.K., Kannu, P., Licht, C., and Carter, M.T. (2013). The first report of nephrocalcinosis in a patient with a 16q23.1-16q23.3 deletion, global developmental delay, trigonocephaly, and portocaval shunt. *Clin Dysmorphol* 152–155.
- de Jong, E.M., Felix, J.F., de Klein, A., and Tibboel, D. (2010a). Etiology of Esophageal Atresia and Tracheoesophageal Fistula: “Mind the Gap.” *Curr. Gastroenterol. Rep.* *12*, 215–222.
- de Jong, E.M., Douben, H., Eussen, B.H., Felix, J.F., Wessels, M.W., Poddighe, P.J., Nikkels, P.G.J., de Krijger, R.R., Tibboel, D., and de Klein, A. (2010b). 5q11.2 deletion in a patient with tracheal agenesis. *Eur. J. Hum. Genet. EJHG* *18*, 1265–1268.
- Jongmans, M.C.J., Admiraal, R.J., Donk, K.P. van der, Vissers, L.E.L.M., Baas, A.F., Kapusta, L., Hagen, J.M. van, Donnai, D., Ravel, T.J. de, Veltman, J.A., et al. (2006). CHARGE syndrome: the phenotypic spectrum of mutations in the *CHD7* gene. *J. Med. Genet.* *43*, 306–314.
- Kaku, Y., Ohmori, T., Kudo, K., Fujimura, S., Suzuki, K., Evans, S.M., Kawakami, Y., and Nishinakamura, R. (2013). *Islet1* deletion causes kidney agenesis and hydroureter resembling CAKUT. *J. Am. Soc. Nephrol. JASN* *24*, 1242–1249.
- Källén, B., Finnström, O., Nygren, K.G., and Olausson, P.O. (2005). In vitro fertilization (IVF) in Sweden: risk for congenital malformations after different IVF methods. *Birt. Defects Res. A. Clin. Mol. Teratol.* *73*, 162–169.
- Kan, Y.W., and Dozy, A.M. (1978). Polymorphism of DNA sequence adjacent to human beta-globin structural gene: relationship to sickle mutation. *Proc. Natl. Acad. Sci. U. S. A.* *75*, 5631–5635.
- Kang, S., Graham, J.M., Olney, A.H., and Biesecker, L.G. (1997). *GLI3* frameshift mutations cause autosomal dominant Pallister-Hall syndrome. *Nat. Genet.* *15*, 266–268.
- Kawalia, A., Motameny, S., Wonzak, S., Thiele, H., Nieroda, L., Jabbari, K., Borowski, S., Sinha, V., Gunia, W., Lang, U., et al. (2015). Leveraging the Power of High Performance Computing for Next Generation Sequencing Data Analysis: Tricks and Twists from a High Throughput Exome Workflow. *PLOS ONE* *10*, e0126321.
- Kent, W.J., Sugnet, C.W., Furey, T.S., Roskin, K.M., Pringle, T.H., Zahler, A.M., and Haussler, and D. (2002). The Human Genome Browser at UCSC. *Genome Res.* *12*, 996–1006.
- Keppler-Noreuil, K.M. (2001). OEIS complex (omphalocele-exstrophy-imperforate anus-spinal defects): a review of 14 cases. *Am. J. Med. Genet.* *99*, 271–279.
- Kim, J., Kim, P., and Hui, C.-C. (2001). The VACTERL association: lessons from the Sonic hedgehog pathway. *Clin. Genet.* *59*, 306–315.
- Kim, S.M., Chang, H.K., Lee, M.J., Shim, K.W., Oh, J.T., Kim, D.S., Kim, M.J., and Han, S.J. (2010). Spinal dysraphism with anorectal malformation: lumbosacral magnetic resonance imaging evaluation of 120 patients. *J. Pediatr. Surg.* *45*, 769–776.

- Kimmel, C.B., Ballard, W.W., Kimmel, S.R., Ullmann, B., and Schilling, T.F. (1995). Stages of embryonic development of the zebrafish. *Dev. Dyn.* 203, 253–310.
- Kohlhase, J., Wischermann, A., Reichenbach, H., Froster, U., and Engel, W. (1998). Mutations in the SALL1 putative transcription factor gene cause Townes-Brocks syndrome. *Nat. Genet.* 18, 81–83.
- Kolde, R. (2015). Package “pheatmap”. <https://cran.r-project.org/web/packages/pheatmap/index.html>.
- Kosaki, R., Fukuhara, Y., Kosuga, M., Okuyama, T., Kawashima, N., Honna, T., Ueoka, K., and Kosaki, K. (2005). OEIS complex with del(3)(q12.2q13.2). *Am. J. Med. Genet. A.* 135, 224–226.
- Kraft, P., Zeggini, E., and Ioannidis, J.P.A. (2009). Replication in genome-wide association studies. *Stat. Sci. Rev. J. Inst. Math. Stat.* 24, 561–573.
- Krepischi, A.C.V., Pearson, P.L., and Rosenberg, C. (2012). Germline copy number variations and cancer predisposition. *Future Oncol. Lond. Engl.* 8, 441–450.
- Kroeger, P.T., and Wingert, R.A. (2014). Using Zebrafish to Study Podocyte Genesis During Kidney Development and Regeneration. *Genes. N. Y. N 2000* 52, 771–792.
- Kuhlenbäumer, G., Hullmann, J., and Appenzeller, S. (2011). Novel genomic techniques open new avenues in the analysis of monogenic disorders. *Hum. Mutat.* 32, 144–151.
- Kyyrönen, P., and Hemminki, K. (1988). Gastro-Intestinal Atresias in Finland in 1970-79, Indicating Time-Place Clustering. *J. Epidemiol. Community Health* 1979- 42, 257–265.
- Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., et al. (2001). Initial sequencing and analysis of the human genome. *Nature* 409, 860–921.
- Lee, C., Reutter, H.M., Grässer, M.F., Fisch, M., and Noeker, M. (2006). Gender-associated differences in the psychosocial and developmental outcome in patients affected with the bladder exstrophy-epispadias complex. *BJU Int.* 97, 349–353.
- Lek, M., Karczewski, K.J., Minikel, E.V., Samocha, K.E., Banks, E., Fennell, T., O’Donnell-Luria, A.H., Ware, J.S., Hill, A.J., Cummings, B.B., et al. (2016). Analysis of protein-coding genetic variation in 60,706 humans. *Nature* 536, 285–291.
- Levitt, M.A., and Peña, A. (2007). Anorectal malformations. *Orphanet J. Rare Dis.* 2, 33.
- Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* 25, 1754–1760.
- Liang, Y., Shen, D., and Cai, W. (2008). Two coding single nucleotide polymorphisms in the SALL1 gene in Townes-Brocks syndrome: a case report and review of the literature. *J. Pediatr. Surg.* 43, 391–393.
- Liao, Y., Smyth, G.K., and Shi, W. (2013). The Subread aligner: fast, accurate and scalable read mapping by seed-and-vote. *Nucleic Acids Res.* 41, e108.
- Liao, Y., Smyth, G.K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinforma. Oxf. Engl.* 30, 923–930.

- Liu, X., Jian, X., and Boerwinkle, E. (2011). dbNSFP: A lightweight database of human nonsynonymous SNPs and their functional predictions. *Hum. Mutat.* *32*, 894–899.
- Liu, X., Wu, C., Li, C., and Boerwinkle, E. (2016). dbNSFP v3.0: A One-Stop Database of Functional Predictions and Annotations for Human Nonsynonymous and Splice-Site SNVs. *Hum. Mutat.* *37*, 235–241.
- Loane, M., Dolk, H., Morris, J.K., and EUROCAT Working Group (2009). Maternal age-specific risk of non-chromosomal anomalies. *BJOG Int. J. Obstet. Gynaecol.* *116*, 1111–1119.
- von Lowtzow, C., Hofmann, A., Zhang, R., Marsch, F., Ebert, A.-K., Rösch, W., Stein, R., Boemers, T.M., Hirsch, K., Marcelis, C., et al. (2016). CNV analysis in 169 patients with bladder exstrophy-epispadias complex. *BMC Med. Genet.* *17*, 35.
- Ludwig, M., Ching, B., Reutter, H., and Boyadjiev, S.A. (2009). Bladder exstrophy-epispadias complex. *Birt. Defects Res. A. Clin. Mol. Teratol.* *85*, 509–522.
- Lundin, J., Söderhäll, C., Lundén, L., Hammarsjö, A., White, I., Schoumans, J., Läckgren, G., Kockum, C.C., and Nordenskjöld, A. (2010). 22q11.2 microduplication in two patients with bladder exstrophy and hearing impairment. *Eur. J. Med. Genet.* *53*, 61–65.
- Maas, S.M., Shaw, A.C., Bikker, H., Lüdecke, H.-J., van der Tuin, K., Badura-Stronka, M., Belligni, E., Biamino, E., Bonati, M.T., Carvalho, D.R., et al. (2015). Phenotype and genotype in 103 patients with tricho-rhino-phalangeal syndrome. *Eur. J. Med. Genet.* *58*, 279–292.
- MacArthur, D.G., Manolio, T.A., Dimmock, D.P., Rehm, H.L., Shendure, J., Abecasis, G.R., Adams, D.R., Altman, R.B., Antonarakis, S.E., Ashley, E.A., et al. (2014). Guidelines for investigating causality of sequence variants in human disease. *Nature* *508*, 469–476.
- Macatee, T.L., Hammond, B.P., Arenkiel, B.R., Francis, L., Frank, D.U., and Moon, A.M. (2003). Ablation of specific expression domains reveals discrete functions of ectoderm- and endoderm-derived FGF8 during cardiovascular and pharyngeal development. *Dev. Camb. Engl.* *130*, 6361–6374.
- Majewski, J., Schwartzenruber, J., Lalonde, E., Montpetit, A., and Jabado, N. (2011). What can exome sequencing do for you? *J. Med. Genet.* *48*, 580–589.
- Männer, J., and Kluth, D. (2003). A chicken model to study the embryology of cloacal exstrophy. *J. Pediatr. Surg.* *38*, 678–681.
- Marra, A.N., and Wingert, R.A. (2014). Roles of Iroquois Transcription Factors in Kidney Development. *Cell Dev. Biol.* *3*, 1000131.
- Martínez-Fernández, M., Rubio, C., Segovia, C., López-Calderón, F.F., Dueñas, M., and Paramio, J.M. (2015). EZH2 in Bladder Cancer, a Promising Therapeutic Target. *Int. J. Mol. Sci.* *16*, 27107–27132.
- Martínez-Frías, M.L., Bermejo, E., Rodríguez-Pinilla, E., and Frías, J.L. (2001). Exstrophy of the cloaca and exstrophy of the bladder: two different expressions of a primary developmental field defect. *Am. J. Med. Genet.* *99*, 261–269.
- Mastroiacovo, P., Castilla, E.E., Arpino, C., Botting, B., Cocchi, G., Goujard, J., Marinacci, C., Merlob, P., Métneki, J., Mutchinick, O., et al. (1999). Congenital malformations in twins: An international study. *Am. J. Med. Genet.* *83*, 117–124.

- McCarthy, D.J., and Smyth, G.K. (2009). Testing significance relative to a fold-change threshold is a TREAT. *Bioinformatics* 25, 765–771.
- McCarthy, M.I., Abecasis, G.R., Cardon, L.R., Goldstein, D.B., Little, J., Ioannidis, J.P.A., and Hirschhorn, J.N. (2008). Genome-wide association studies for complex traits: consensus, uncertainty and challenges. *Nat. Rev. Genet.* 9, 356–369.
- McDonald-McGinn, D.M., Sullivan, K.E., Marino, B., Philip, N., Swillen, A., Vorstman, J.A.S., Zackai, E.H., Emanuel, B.S., Vermeesch, J.R., Morrow, B.E., et al. (2015). 22q11.2 deletion syndrome. *Nat. Rev. Dis. Primer* 1, 15071.
- McHONEY, M., RANSLEY, P., CUCKOW, P., KIELY, E., and SPITZ, L. (2001). DUODENAL ATRESIA ASSOCIATED WITH MIDGUT DELETION IN CLOACAL EXSTROPHY: A NEW ASSOCIATION? *J. Urol.* 166, 1041.
- McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., et al. (2010). The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 20, 1297–1303.
- McNeal, R.M., Skoglund, R.R., and Francke, U. (1977). Congenital anomalies including the VATER association in a patient with a del(6)q deletion. *J. Pediatr.* 91, 957–960.
- Midrio, P., Nogare, C.D., Di Gianantonio, E., and Clementi, M. (2006). Are congenital anorectal malformations more frequent in newborns conceived with assisted reproductive techniques? *Reprod. Toxicol. Elmsford N* 22, 576–577.
- Miller, E.A., Manning, S.E., Rasmussen, S.A., Reefhuis, J., Honein, M.A., and National Birth Defects Prevention Study (2009). Maternal exposure to tobacco smoke, alcohol and caffeine, and risk of anorectal atresia: National Birth Defects Prevention Study 1997-2003. *Paediatr. Perinat. Epidemiol.* 23, 9–17.
- Mills, R.E., Walter, K., Stewart, C., Handsaker, R.E., Chen, K., Alkan, C., Abyzov, A., Yoon, S.C., Ye, K., Cheetham, R.K., et al. (2011). Mapping copy number variation by population-scale genome sequencing. *Nature* 470, 59–65.
- Mittal, A., Airon, R.K., Magu, S., Rattan, K.N., and Ratan, S.K. (2004). Associated anomalies with anorectal malformation (ARM). *Indian J. Pediatr.* 71, 509–514.
- Motoyama, J., Liu, J., Mo, R., Ding, Q., Post, M., and Hui, C. (1998). Essential function of Gli2 and Gli3 in the formation of lung, trachea and oesophagus. *Nat. Genet.* 20, 54–57.
- Moura, E., Cirio, S.M., and Villanova, J.A. (2010). VACTERL association in a cat. *Am. J. Med. Genet. A.* 152A, 777–780.
- Muecke, E.C. (1964). THE ROLE OF THE CLOACAL MEMBRANE IN EXSTROPHY: THE FIRST SUCCESSFUL EXPERIMENTAL STUDY. *J. Urol.* 92, 659–667.
- Myers, M.F., Li, S., Correa-Villaseñor, A., Li, Z., Moore, C.A., Hong, S.X., Berry, R.J., and China-US Collaborative Project for Neural Tube Defect Prevention (2001). Folic acid supplementation and risk for imperforate anus in China. *Am. J. Epidemiol.* 154, 1051–1056.
- Nakamura, Y., Tsiaris, C.D., Özbek, S., and Holstein, T.W. (2011). Autoregulatory and repressive inputs localize Hydra Wnt3 to the head organizer. *Proc. Natl. Acad. Sci. U. S. A.* 108, 9137–9142.

- Nassar, N., Leoncini, E., Amar, E., Arteaga-Vázquez, J., Bakker, M.K., Bower, C., Canfield, M.A., Castilla, E.E., Cocchi, G., Correa, A., et al. (2012). Prevalence of Esophageal Atresia among 18 International Birth Defects Surveillance Programs. *Birt. Defects Res. A. Clin. Mol. Teratol.* *94*, 893–899.
- Ngan, E.S.-W., Kim, K.-H., and Hui, C. -c. (2013). Sonic Hedgehog Signaling and VACTERL Association. *Mol. Syndromol.* *4*, 32–45.
- Niemi, J., Mittman, E., Landau, W., and Nettleton, D. (2015). Empirical Bayes analysis of RNA-seq data for detection of gene expression heterosis. *J. Agric. Biol. Environ. Stat.* *20*, 614–628.
- Niegelstein, R.A., Valk, J., Smit, L.M., and Vermeij-Keers, C. (1994). MR of the caudal regression syndrome: embryologic implications. *AJNR Am. J. Neuroradiol.* *15*, 1021–1029.
- Niegelstein, R.A., van der Werff, J.F., Verbeek, F.J., Valk, J., and Vermeij-Keers, C. (1998). Normal and abnormal embryonic development of the anorectum in human embryos. *Teratology* *57*, 70–78.
- Niegelstein, R. a. J., Hartwig, N.G., Vermeij-Keers, C., and Valk, J. (1993). Embryonic development of the mammalian caudal Neural Tube. *Teratology* *48*, 21–31.
- Niegelstein, R. a. J., Vos, A., Valk, J., and Vermeij-Keers, C. (2002). Magnetic resonance imaging in children with anorectal malformations: Embryologic implications. *J. Pediatr. Surg.* *37*, 1138–1145.
- Noël, A.-C., Pelluard, F., Delezoide, A.-L., Devisme, L., Loeuillet, L., Leroy, B., Martin, A., Bouvier, R., Laquerriere, A., Jeanne-Pasquier, C., et al. (2014). Fetal phenotype associated with the 22q11 deletion. *Am. J. Med. Genet. A.* *164*, 2724–2731.
- Nora, J.J., Nora, A.H., Blu, J., Ingram, J., Fountain, A., Peterson, M., Lortscher, R.H., and Kimberling, W.J. (1978). Exogenous Progestogen and Estrogen Implicated in Birth Defects. *JAMA* *240*, 837–843.
- Nutsua, M.E., Fischer, A., Nebel, A., Hofmann, S., Schreiber, S., Krawczak, M., and Nothnagel, M. (2015). Family-Based Benchmarking of Copy Number Variation Detection Software. *PLOS ONE* *10*, e0133465.
- Oeth, P., Beaulieu, M., Park, C., Kosman, D., del Mistro, G., van den Boom, D., Jurinke, C., Paul Oeth, M.B., and Jurinke, C. (2005). iPLEX™ assay: increased plexing efficiency and flexibility for MassARRAY® system through single base primer extension with mass-modified terminators.
- Ogawa, K., Utsunomiya, T., Mimori, K., Tanaka, Y., Tanaka, F., Inoue, H., Murayama, S., and Mori, M. (2004). Clinical significance of elongation factor-1 delta mRNA expression in oesophageal carcinoma. *Br. J. Cancer* *91*, 282–286.
- Peddibhotla, S., Khalifa, M., Probst, F.J., Stein, J., Harris, L.L., Kearney, D.L., Vance, G.H., Bull, M.J., Grange, D.K., Scharer, G.H., et al. (2013). Expanding the genotype-phenotype correlation in subtelomeric 19p13.3 microdeletions using high resolution clinical chromosomal microarray analysis. *Am. J. Med. Genet. A.* *161A*, 2953–2963.
- Peiffer, D.A., Le, J.M., Steemers, F.J., Chang, W., Jenniges, T., Garcia, F., Haden, K., Li, J., Shaw, C.A., Belmont, J., et al. (2006). High-resolution genomic profiling of chromosomal aberrations using Infinium whole-genome genotyping. *Genome Res.* *16*, 1136–1148.
- Petersen, B.-S., Fredrich, B., Hoepfner, M.P., Ellinghaus, D., and Franke, A. (2017). Opportunities and challenges of whole-genome and -exome sequencing. *BMC Genet.* *18*, 14.

- Pinheiro, P.F.M., Simões e Silva, A.C., and Pereira, R.M. (2012). Current knowledge on esophageal atresia. *World J. Gastroenterol. WJG* 18, 3662–3672.
- Prieto, J.C., Garcia, N.M., Elder, F.F., Zinn, A.R., and Baker, L.A. (2007). Phenotypic expansion of the supernumerary derivative (22) chromosome syndrome: VACTERL and Hirschsprung's disease. *J. Pediatr. Surg.* 42, 1928–1932.
- van der Putte, S.C. (1986). Normal and abnormal development of the anorectum. *J. Pediatr. Surg.* 21, 434–440.
- van der Putte, S.C.J., Spliet, W.G.M., and Nikkels, P.G.J. (2008). Common (“classical”) and covered cloacal exstrophy: a histopathological study and a reconstruction of the pathogenesis. *Pediatr. Dev. Pathol. Off. J. Soc. Pediatr. Pathol. Paediatr. Pathol. Soc.* 11, 430–442.
- Pyati, U.J., Cooper, M.S., Davidson, A.J., Nechiporuk, A., and Kimelman, D. (2006). Sustained Bmp signaling is essential for cloaca development in zebrafish. *Dev. Camb. Engl.* 133, 2275–2284.
- Qi, L., Wang, M., Yagnik, G., Mattheisen, M., Gearhart, J.P., Lakshmanan, Y., Ebert, A.-K., Rösch, W., Ludwig, M., Draaken, M., et al. (2013). Candidate gene association study implicates p63 in the etiology of nonsyndromic bladder-exstrophy-epispadias complex. *Birt. Defects Res. A. Clin. Mol. Teratol.* 97, 759–763.
- Quan, L., and Smith, D.W. (1973). The VATER association. Vertebral defects, Anal atresia, T-E fistula with esophageal atresia, Radial and Renal dysplasia: a spectrum of associated defects. *J. Pediatr.* 82, 104–107.
- Ramu, A., Noordam, M.J., Schwartz, R.S., Wuster, A., Hurles, M.E., Cartwright, R.A., and Conrad, D.F. (2013). DeNovoGear: de novo indel and point mutation discovery and phasing. *Nat. Methods* 10, 985–987.
- Reardon, W., Zhou, X.P., and Eng, C. (2001). A novel germline mutation of the PTEN gene in a patient with macrocephaly, ventricular dilatation, and features of VATER association. *J. Med. Genet.* 38, 820–823.
- Redon, R., Ishikawa, S., Fitch, K.R., Feuk, L., Perry, G.H., Andrews, T.D., Fiegler, H., Shapero, M.H., Carson, A.R., Chen, W., et al. (2006). Global variation in copy number in the human genome. *Nature* 444, 444–454.
- Reefhuis, J., Honein, M.A., Schieve, L.A., Correa, A., Hobbs, C.A., Rasmussen, S.A., and National Birth Defects Prevention Study (2009). Assisted reproductive technology and major structural birth defects in the United States. *Hum. Reprod. Oxf. Engl.* 24, 360–366.
- Reefhuis, J., Honein, M.A., Schieve, L.A., Rasmussen, S.A., and National Birth Defects Prevention Study (2011). Use of clomiphene citrate and birth defects, National Birth Defects Prevention Study, 1997–2005. *Hum. Reprod. Oxf. Engl.* 26, 451–457.
- Reutter, H., and Ludwig, M. (2013). VATER/VACTERL Association: Evidence for the Role of Genetic Factors. *Mol. Syndromol.* 4, 16–19.
- Reutter, H., Shapiro, E., and Gruen, J.R. (2003). Seven new cases of familial isolated bladder exstrophy and epispadias complex (BEEC) and review of the literature. *Am. J. Med. Genet. A.* 120A, 215–221.

- Reutter, H., Becker, T., Ludwig, M., Schäfer, N., Detlefsen, B., Beaudoin, S., Fisch, M., Ebert, A.-K., Rösch, W., Nöthen, M.M., et al. (2006). Family-based association study of the MTHFR polymorphism C677T in the bladder-exstrophy-epispadias-complex. *Am. J. Med. Genet. A.* *140*, 2506–2509.
- Reutter, H., Qi, L., Gearhart, J.P., Boemers, T., Ebert, A.-K., Rösch, W., Ludwig, M., and Boyadjiev, S.A. (2007). Concordance analyses of twins with bladder exstrophy-epispadias complex suggest genetic etiology. *Am. J. Med. Genet. A.* *143A*, 2751–2756.
- Reutter, H., Boyadjiev, S.A., Gambhir, L., Ebert, A.-K., Rösch, W.H., Stein, R., Schröder, A., Boemers, T.M., Bartels, E., Vogt, H., et al. (2011). Phenotype severity in the bladder exstrophy-epispadias complex: analysis of genetic- and non-genetic contributing factors in 441 families from North America and Europe. *J. Pediatr.* *159*, 825–831.e1.
- Reutter, H., Gurung, N., and Ludwig, M. (2014a). Evidence for annular pancreas as an associated anomaly in the VATER/VACTERL association and investigation of the gene encoding pancreas specific transcription factor 1A as a candidate gene. *Am. J. Med. Genet. A.* *164*, 1611–1613.
- Reutter, H., Draaken, M., Pennimpede, T., Wittler, L., Brockschmidt, F.F., Ebert, A.-K., Bartels, E., Rösch, W., Boemers, T.M., Hirsch, K., et al. (2014b). Genome-wide association study and mouse expression data identify a highly conserved 32 kb intergenic region between WNT3 and WNT9b as possible susceptibility locus for isolated classic exstrophy of the bladder. *Hum. Mol. Genet.* *23*, 5536–5544.
- Reutter, H., Hilger, A.C., Hildebrandt, F., and Ludwig, M. (2016a). Underlying genetic factors of the VATER/VACTERL association with special emphasis on the “Renal” phenotype. *Pediatr. Nephrol. Berl. Ger.* *31*, 2025–2033.
- Reutter, H., Keppler-Noreuil, K., E. Keegan, C., Thiele, H., Yamada, G., and Ludwig, M. (2016b). Genetics of Bladder-Exstrophy-Epispadias Complex (BEEC): Systematic Elucidation of Mendelian and Multifactorial Phenotypes. *Curr. Genomics* *17*, 4–13.
- Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., Grody, W.W., Hegde, M., Lyon, E., Spector, E., et al. (2015). Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet. Med. Off. J. Am. Coll. Med. Genet.* *17*, 405–424.
- Riis, B., Rattan, S.I., Clark, B.F., and Merrick, W.C. (1990). Eukaryotic protein elongation factors. *Trends Biochem. Sci.* *15*, 420–424.
- Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W., and Smyth, G.K. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* *43*, e47.
- Robert, E., Mutchinick, O., Mastroiacovo, P., Knudsen, L.B., Daltveit, A.K., Castilla, E.E., Lancaster, P., Källén, B., and Cocchi, G. (1993). An international collaborative study of the epidemiology of esophageal atresia or stenosis. *Reprod. Toxicol. Elmsford N* *7*, 405–421.
- Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinforma. Oxf. Engl.* *26*, 139–140.
- Rovelet-Lecrux, A., Hannequin, D., Raux, G., Meur, N.L., Laquerrière, A., Vital, A., Dumanchin, C., Feuillette, S., Brice, A., Vercelletto, M., et al. (2006). APP locus duplication causes autosomal dominant early-onset Alzheimer disease with cerebral amyloid angiopathy. *Nat. Genet.* *38*, 24–26.

- Saisawat, P., Kohl, S., Hilger, A.C., Hwang, D.-Y., Yung Gee, H., Dworschak, G.C., Tasic, V., Pennimpede, T., Natarajan, S., Sperry, E., et al. (2014). Whole-exome resequencing reveals recessive mutations in TRAP1 in individuals with CAKUT and VACTERL association. *Kidney Int.* *85*, 1310–1317.
- Sajan, S.A., Fernandez, L., Nieh, S.E., Rider, E., Bukshpun, P., Wakahiro, M., Christian, S.L., Rivière, J.-B., Sullivan, C.T., Sudi, J., et al. (2013). Both Rare and De Novo Copy Number Variants Are Prevalent in Agenesis of the Corpus Callosum but Not in Cerebellar Hypoplasia or Polymicrogyria. *PLoS Genet.* *9*.
- Sanger, F., Nicklen, S., and Coulson, A.R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. U. S. A.* *74*, 5463–5467.
- Schmermund, A., Möhlenkamp, S., Stang, A., Grönemeyer, D., Seibel, R., Hirche, H., Mann, K., Siffert, W., Lauterbach, K., Siegrist, J., et al. (2002). Assessment of clinically silent atherosclerotic disease and established and novel risk factors for predicting myocardial infarction and cardiac death in healthy middle-aged subjects: rationale and design of the Heinz Nixdorf RECALL Study. *Risk Factors, Evaluation of Coronary Calcium and Lifestyle. Am. Heart J.* *144*, 212–218.
- Schnitzer, P.G., Olshan, A.F., and Erickson, J.D. (1995). Paternal occupation and risk of birth defects in offspring. *Epidemiol. Camb. Mass* *6*, 577–583.
- Schramm, C., Draaken, M., Bartels, E., Boemers, T.M., Aretz, S., Brockschmidt, F.F., Nöthen, M.M., Ludwig, M., and Reutter, H. (2011). De novo microduplication at 22q11.21 in a patient with VACTERL association. *Eur. J. Med. Genet.* *54*, 9–13.
- Scott, D.A. (1993). Esophageal Atresia/Tracheoesophageal Fistula Overview. In *GeneReviews*(®), R.A. Pagon, M.P. Adam, H.H. Ardinger, S.E. Wallace, A. Amemiya, L.J. Bean, T.D. Bird, N. Ledbetter, H.C. Mefford, R.J. Smith, et al., eds. (Seattle (WA): University of Washington, Seattle), p.
- Seo, J., Kim, D.Y., Kim, A.R., Kim, D.Y., Kim, S.C., Kim, I.K., Kim, K.S., Yoon, C.H., and Pi, S.Y. (2010). An 18-year experience of tracheoesophageal fistula and esophageal atresia. *Korean J. Pediatr.* *53*, 705–710.
- Shaw-Smith, C. (2006). Oesophageal atresia, tracheo-oesophageal fistula, and the VACTERL association: review of genetics and epidemiology. *J. Med. Genet.* *43*, 545–554.
- Shaw-Smith, C. (2010). Genetic factors in esophageal atresia, tracheo-esophageal fistula and the VACTERL association: Roles for FOXF1 and the 16q24.1 FOX transcription factor gene cluster, and review of the literature. *Eur. J. Med. Genet.* *53*, 6–13.
- Singleton, A.B., Farrer, M., Johnson, J., Singleton, A., Hague, S., Kachergus, J., Hulihan, M., Peuralinna, T., Dutra, A., Nussbaum, R., et al. (2003). α -Synuclein Locus Triplication Causes Parkinson9s Disease. *Science* *302*, 841–841.
- Smigiel, R., Marcelis, C., Patkowski, D., de Leeuw, N., Bednarczyk, D., Barg, E., Mascianica, K., Maria Sasiadek, M., and Brunner, H. (2014). Oesophageal atresia with tracheoesophageal fistula and anal atresia in a patient with a de novo microduplication in 17q12. *Eur. J. Med. Genet.* *57*, 40–43.
- Smyth, G.K. (2004). Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat. Appl. Genet. Mol. Biol.* *3*, Article3.
- Solomon, B.D., Pineda–Alvarez, D.E., Hadley, D.W., Keaton, A.A., Agochukwu, N.B., Raam, M.S., Carlson–Donohoe, H.E., Kamat, A., and Chandrasekharappa, S.C. (2011). De Novo Deletion of Chromosome 20q13.33 in a Patient with Tracheo-esophageal Fistula, Cardiac Defects and

- Genitourinary Anomalies Implicates GTPBP5 as a Candidate Gene. *Birt. Defects Res. A. Clin. Mol. Teratol.* *91*, 862–865.
- Solomon, B.D., Pineda-Alvarez, D.E., Hadley, D.W., Hansen, N.F., Kamat, A., Donovan, F.X., Chandrasekharappa, S.C., Hong, S.-K., Roessler, E., and Mullikin, J.C. (2013). Exome Sequencing and High-Density Microarray Testing in Monozygotic Twin Pairs Discordant for Features of VACTERL Association. *Mol. Syndromol.* *4*, 27–31.
- Solomon, B.D., Baker, L.A., Bear, K.A., Cunningham, B.K., Giampietro, P.F., Hadigan, C., Hadley, D.W., Harrison, S., Levitt, M.A., Niforatos, N., et al. (2014). An Approach to the Identification of Anomalies and Etiologies in Neonates with Identified or Suspected VACTERL (Vertebral Defects, Anal Atresia, Tracheo-Esophageal Fistula with Esophageal Atresia, Cardiac Anomalies, Renal Anomalies, and Limb Anomalies) Association. *J. Pediatr.* *164*, 451–457.e1.
- Southard, A.E., Edlmann, L.J., and Gelb, B.D. (2012). Role of Copy Number Variants in Structural Birth Defects. *Pediatrics* *129*, 755–763.
- Stankiewicz, P., and Lupski, J.R. (2010). Structural Variation in the Human Genome and its Role in Disease. *Annu. Rev. Med.* *61*, 437–455.
- Stemers, F.J., Chang, W., Lee, G., Barker, D.L., Shen, R., and Gunderson, K.L. (2006). Whole-genome genotyping with the single-base extension assay. *Nat. Methods* *3*, 31–33.
- Steward, C.A., Parker, A.P.J., Minassian, B.A., Sisodiya, S.M., Frankish, A., and Harrow, J. (2017). Genome annotation for clinical genomic diagnostics: strengths and weaknesses. *Genome Med.* *9*, 49.
- Stoll, C., Alembik, Y., Roth, M.P., and Dott, B. (1997). Risk factors in congenital anal atresias. *Ann. Genet.* *40*, 197–204.
- Stoll, C., Alembik, Y., Dott, B., and Roth, M.P. (2007). Associated malformations in patients with anorectal anomalies. *Eur. J. Med. Genet.* *50*, 281–290.
- Sun, J., Liu, Y.-H., Chen, H., Nguyen, M.P., Mishina, Y., Upperman, J.S., Ford, H.R., and Shi, W. (2007). Deficient Alk3-mediated BMP signaling causes prenatal omphalocele-like defect. *Biochem. Biophys. Res. Commun.* *360*, 238–243.
- Szumaska, D., Pielas, G., Essalmani, R., Bilski, M., Mesnard, D., Kaur, K., Franklyn, A., El Omari, K., Jefferis, J., Bentham, J., et al. (2008). VACTERL/caudal regression/Currarino syndrome-like malformations in mice with mutation in the proprotein convertase Pcsk5. *Genes Dev.* *22*, 1465–1477.
- Thauvin-Robinet, C., Faivre, L., Cusin, V., Khau Van Kien, P., Callier, P., Parker, K.L., Fellous, M., Borgnon, J., Gounot, E., Huet, F., et al. (2004). Cloacal exstrophy in an infant with 9q34.1-qter deletion resulting from a de novo unbalanced translocation between chromosome 9q and Yq. *Am. J. Med. Genet. A.* *126A*, 303–307.
- The 1000 Genomes Project Consortium (2010). A map of human genome variation from population-scale sequencing. *Nature* *467*, 1061–1073.
- The 1000 Genomes Project Consortium (2012). An integrated map of genetic variation from 1,092 human genomes. *Nature* *491*, 56–65.
- The 1000 Genomes Project Consortium (2015). A global reference for human genetic variation. *Nature* *526*, 68–74.

- The International HapMap Consortium (2010). Integrating common and rare genetic variation in diverse human populations. *Nature* 467, 52–58.
- Thisse, C., and Thisse, B. (2008). High-resolution in situ hybridization to whole-mount zebrafish embryos. *Nat. Protoc.* 3, 59–69.
- Thomalla, J.V., Rudolph, R.A., Rink, R.C., and Mitchell, M.E. (1985). Induction of cloacal exstrophy in the chick embryo using the CO₂ laser. *J. Urol.* 134, 991–995.
- Tobias, N., Mason, D., Lutkenhoff, M., Stoops, M., and Ferguson, D. (2008). Management Principles of Organic Causes of Childhood Constipation. *J Pediatr Health Care* 22(1):12-23.
- Torfs, C.P., Curry, C.J., and Bateson, T.F. (1995). Population-based study of tracheoesophageal fistula and esophageal atresia. *Teratology* 52, 220–232.
- Tubbs, R.S., Smyth, M.D., and Oakes, W.J. (2003). Chiari I malformation and cloacal exstrophy: report of a patient with both defects of blastogenesis. *Am. J. Med. Genet. A.* 119A, 231–233.
- Uchimura, T., Komatsu, Y., Tanaka, M., McCann, K.L., and Mishina, Y. (2009). Bmp2 and Bmp4 genetically interact to support multiple aspects of mouse development including functional heart development. *Genes. N. Y. N* 2000 47, 374–384.
- Vedel-Macrandar, G.C., and Hood, R.D. (1986). Teratogenic effects of nigericin, a carboxylic ionophore. *Teratology* 33, 47–51.
- van der Veken, L.T., Dieleman, M.M., Douben, H., van de Brug, J.C., van de Graaf, R., Hoogeboom, A.J.M., Poddighe, P.J., and de Klein, A. (2010). Low grade mosaic for a complex supernumerary ring chromosome 18 in an adult patient with multiple congenital anomalies. *Mol. Cytogenet.* 3, 13.
- Vlangos, C.N., Siuniak, A., Ackley, T., van Bokhoven, H., Veltman, J., Iyer, R., Park, J.M., Keppler-Noreuil, K., and Keegan, C.E. (2011). Comprehensive genetic analysis of OEIS complex reveals no evidence for a recurrent microdeletion or duplication. *Am. J. Med. Genet. A.* 155A, 38–49.
- Wallingford, M.C., and Giachelli, C.M. (2014). Loss of PiT-1 results in abnormal endocytosis in the yolk sac visceral endoderm. *Mech. Dev.* 133, 189–202.
- Wei, X., and Sulik, K.K. (1996). Pathogenesis of caudal dysgenesis/sirenomelia induced by ochratoxin A in chick embryos. *Teratology* 53, 378–391.
- Wessels, M.W., Kuchinka, B., Heydanus, R., Smit, B.J., Dooijes, D., de Krijger, R.R., Lequin, M.H., de Jong, E.M., Husen, M., Willems, P.J., et al. (2010). Polyalanine expansion in the ZIC3 gene leading to X-linked heterotaxy with VACTERL association: a new polyalanine disorder? *J. Med. Genet.* 47, 351–355.
- Wijers, C.H.W., Blaauw, I. de, Marcelis, C.L.M., Wijnen, R.M.H., Brunner, H., Midrio, P., Gamba, P., Clementi, M., Jenetzky, E., Zwink, N., et al. (2010). Research perspectives in the etiology of congenital anorectal malformations using data of the International Consortium on Anorectal Malformations: evidence for risk factors across different populations. *Pediatr. Surg. Int.* 26, 1093–1099.
- Wilkins, S., Zhang, K.W., Mahfuz, I., Quantin, R., D’Cruz, N., Hutson, J., Ee, M., Bagli, D., Aitken, K., Fong, F.N.-Y., et al. (2012). Insertion/deletion polymorphisms in the ΔNp63 promoter are a risk factor for bladder exstrophy epispadias complex. *PLoS Genet.* 8, e1003070.

- Williamson, K.A., Hever, A.M., Rainger, J., Rogers, R.C., Magee, A., Fiedler, Z., Keng, W.T., Sharkey, F.H., McGill, N., Hill, C.J., et al. (2006). Mutations in SOX2 cause anophthalmia-esophageal-genital (AEG) syndrome. *Hum. Mol. Genet.* *15*, 1413–1422.
- Wingert, R.A., Selleck, R., Yu, J., Song, H.-D., Chen, Z., Song, A., Zhou, Y., Thisse, B., Thisse, C., McMahon, A.P., et al. (2007). The *cdx* genes and retinoic acid control the positioning and segmentation of the zebrafish pronephros. *PLoS Genet.* *3*, 1922–1938.
- Wood, H.M., Trock, B.J., and Gearhart, J.P. (2003). In vitro fertilization and the cloacal-bladder exstrophy-epispadias complex: is there an association? *J. Urol.* *169*, 1512–1515.
- Wood, H.M., Babineau, D., and Gearhart, J.P. (2007). In vitro fertilization and the cloacal/bladder exstrophy-epispadias complex: a continuing association. *J. Pediatr. Urol.* *3*, 305–310.
- Worthington, S., Colley, A., Fagan, K., Dai, K., and Lipson, A.H. (1997). Anal anomalies: an uncommon feature of velocardiofacial (Shprintzen) syndrome? *J. Med. Genet.* *34*, 79–82.
- Yamada, K., Uchiyama, A., Arai, M., Kubodera, K., Yamamoto, Y., Orii, K.O., Nagasawa, H., Masuno, M., and Kohno, Y. (2009). Severe upper airway stenosis in a boy with partial monosomy 16p13.3pter and partial trisomy 16q22qter. *Congenit. Anom.* *49*, 85–88.
- Yan, J., Zhang, L., Xu, J., Sultana, N., Hu, J., Cai, X., Li, J., Xu, P.-X., and Cai, C.-L. (2014). Smad4 regulates ureteral smooth muscle cell differentiation during mouse embryogenesis. *PLoS One* *9*, e104503.
- Yeo, G., and Burge, C.B. (2004). Maximum Entropy Modeling of Short Sequence Motifs with Applications to RNA Splicing Signals. *J. Comput. Biol.* *11*, 377–394.
- Yuan, P., Okazaki, I., and Kuroki, Y. (1995). Anal atresia: effect of smoking and drinking habits during pregnancy. *Jpn. J. Hum. Genet.* *40*, 327–332.
- Zeidler, C., Woelfle, J., Draaken, M., Mughal, S.S., Große, G., Hilger, A.C., Dworschak, G.C., Boemers, T.M., Jenetzky, E., Zwink, N., et al. (2014). Heterozygous FGF8 mutations in patients presenting cryptorchidism and multiple VATER/VACTERL features without limb anomalies. *Birt. Defects Res. A. Clin. Mol. Teratol.* *100*, 750–759.
- Zen, P.R.G., Riegel, M., Rosa, R.F.M., Pinto, L.L.C., Graziadio, C., Schwartz, I.V.D., and Paskulin, G.A. (2010). Esophageal stenosis in a child presenting a de novo 7q terminal deletion. *Eur. J. Med. Genet.* *53*, 333–336.
- Zentner, G.E., Layman, W.S., Martin, D.M., and Scacheri, P.C. (2010). Molecular and phenotypic aspects of CHD7 mutation in CHARGE syndrome. *Am. J. Med. Genet. A.* *152A*, 674–686.
- Zhang, F., Gu, W., Hurles, M.E., and Lupski, J.R. (2009). Copy Number Variation in Human Health, Disease, and Evolution. *Annu. Rev. Genomics Hum. Genet.* *10*, 451–481.
- Zhang, R., Thiele, H., Bartmann, P., Hilger, A.C., Berg, C., Herberg, U., Klingmüller, D., Nürnberg, P., Ludwig, M., and Reutter, H. (2016). Whole-Exome Sequencing in Nine Monozygotic Discordant Twins. *Twin Res. Hum. Genet.* *19*, 60–65.
- Zhang, R., Knapp, M., Suzuki, K., Kajioaka, D., Schmidt, J.M., Winkler, J., Yilmaz, Ö., Pleschka, M., Cao, J., Kockum, C.C., et al. (2017a). *ISL1* is a major susceptibility gene for classic bladder exstrophy and a regulator of urinary tract development. *Sci. Rep.* *7*, srep42170.

Zhang, R., Marsch, F., Kause, F., Degenhardt, F., Schmiedeke, E., Märzheuser, S., Hoppe, B., Bachour, H., Boemers, T.M., Schäfer, M., et al. (2017b). Array-based molecular karyotyping in 115 VATER/VACTERL and VATER/VACTERL-like patients identifies disease-causing copy number variations. *Birth Defects Res.* n/a-n/a.

Zon, L.I. (1999). Zebrafish: a new model for human disease. *Genome Res.* 9, 99–100.

Zweier, C., Sticht, H., Aydin-Yaylagül, I., Campbell, C.E., and Rauch, A. (2007). Human TBX1 Missense Mutations Cause Gain of Function Resulting in the Same Phenotype as 22q11.2 Deletions. *Am. J. Hum. Genet.* 80, 510–517.

Zwink, N., Jenetzky, E., Hirsch, K., Reifferscheid, P., Schmiedeke, E., Schmidt, D., Reckin, S., Obermayr, F., Boemers, T.M., Stein, R., et al. (2013). Assisted reproductive techniques and risk of exstrophy-epispadias complex: a German case-control study. *J. Urol.* 189, 1524–1529.

Eigene Publikationen

Weitensteiner, V. ^{*}, **Zhang, R.** ^{*}, Bungenberg, J., Marks, M., Gehlen, J., Ralser, D.J., Hilger, A.C., Sharma, A., Schumacher, J., Gembruch, U., et al. (2018). Exome sequencing in syndromic brain malformations identifies novel mutations in ACTB, and SLC9A6, and suggests BAZ1A as a new candidate gene. *Birth Defects Res.* (* contributed equally)

Kause, F., Reutter, H., Marsch, F., Thiele, H., Altmüller, J., Ludwig, M., and **Zhang, R.** (2018). Whole exome sequencing identifies a mutation in EYA1 and GLI3 in a patient with branchio-otic syndrome and esophageal atresia: Coincidence or a digenic mode of inheritance? *Mol. Med. Rep.* 17, 3200–3205.

Nauman, N., Jalali, S., Shami, S., Rafiq, S., Große, G., Hilger, A.C., **Zhang, R.**, Mansoor, S., Ludwig, M., and Reutter, H. (2018). Low maternal folate concentrations and maternal MTHFR C677T polymorphism are associated with an increased risk for neural tube defects in offspring: a case-control study among Pakistani case and control mothers. *Asia Pac. J. Clin. Nutr.* 27, 253–260.

Zhang, R., Thiele, H., Bartmann, P., Hilger, A.C., Berg, C., Herberg, U., Klingmüller, D., Nürnberg, P., Ludwig, M., and Reutter, H. (2016). Whole-Exome Sequencing in Nine Monozygotic Discordant Twins. *Twin Res. Hum. Genet.* 19, 60–65.

Zhang, R. ^{*}, Knapp, M. ^{*}, Suzuki, K., Kajioka, D., Schmidt, J.M., Winkler, J., Yilmaz, Ö., Pleschka, M., Cao, J., Kockum, C.C., et al. (2017). ISL1 is a major susceptibility gene for classic bladder exstrophy and a regulator of urinary tract development. *Sci. Rep.* 7, srep42170. (* contributed equally)

Zhang, R., Knapp, M., Kause, F., Reutter, H., Ludwig, M., Husman, D.A., Vandersteen, D.R., Hurst, J.A., Gearhart, J.P., Jeffs, R.D., et al. (2017). Role of the LF-SINE–Derived Distal ISL1 Enhancer in Patients with Classic Bladder Exstrophy, Anatomy of the cloacal exstrophy. *J Pediatr Genet.*

Zhang, R., Marsch, F., Kause, F., Degenhardt, F., Schmiedeke, E., Märzheuser, S., Hoppe, B., Bachour, H., Boemers, T.M., Schäfer, M., et al. (2017). Array-based molecular karyotyping in 115 VATER/VACTERL and VATER/VACTERL-like patients identifies disease-causing copy number variations. *Birth Defects Res.* n/a-n/a.

Kolarova, J., Bens, S., Ammerpohl, O., Hilger, A.C., **Zhang, R.**, Reutter, H., and Siebert, R. (2016). PLAGL1 epimutation and bladder exstrophy: Coincidence or concurrent etiology? *Birt. Defects Res. A. Clin. Mol. Teratol.* 106, 724–728.

von Lowtzow, C., Hofmann, A., **Zhang, R.**, Marsch, F., Ebert, A.-K., Rösch, W., Stein, R., Boemers, T.M., Hirsch, K., Marcelis, C., et al. (2016). CNV analysis in 169 patients with bladder exstrophy-epispadias complex. *BMC Med. Genet.* 17, 35.

Appendix A: Material und Methoden

Appendix A1: Alphabetische Auflistung der in dieser Arbeit verwendeten Geräten, Chemikalien, Enzymen, Lösungen, kommerziellen Kits sowie Software und Datenbanken

Geräten

Autoklav

- Systec D-150, Sytec GmbH
- Varioklav® 135 S Dampfsterilisator, H+P Labortechnik GmbH

Automatischer Sequenzier

- 3130xl Genetic Analyzer, Life Technologies Corporation

DNA-Bank/Verwaltung

- 2D CYPHER Tubes, Thermo Fisher Scientific GmbH
- 2D CYPHER 1,2ml Cluster Tube Racks, Thermo Fisher Scientific GmbH
- SmartScan Solo™ 2D Barcode Reader, Thermo Fisher Scientific GmbH
- SmartScan 96 2D Barcode Reader, Thermo Fisher Scientific GmbH
- FluidX 96-well format 2D barcoded sample storage tubes and caps, FluidX GmbH
- X-CAP semi-automatic capmat capper, FluidX GmbH
- FluidX IntelliCode™ Decoding Software, FluidX GmbH

DNA-Extraktion

- Chemagic Magnetic Separation Module I v20, PerkinElmer Chemagen Technologie GmbH

Eismaschine

- AF100, Scotsman® Ice Systems

Gel-Elektrophoreseskammern

- WIDE MINI SUB CELL® GT, BioRad Laboratories GmbH
- Sub-Cell Model 96, BioRad Laboratories GmbH

Gel-Dokumentation

- GelDoc™ XR System, BioRad Laboratories GmbH

Genotypisierungssysteme

- iScan System, Illumina

- Light Cycler®, 480 II, Roche
- MassARRAY™ Compact Analyzer, Bruker Daltonics Inc. für Sequenom® GmbH

Mikrowelle

- MW 9625, Severin

Nanodispenser

- MassARRAY™ Nanodispenser, SAMSUNG Techwin Co. Ltd. Für Sequenom® GmbH

Next-Generation-Sequenzierung

- HiSeq 2500, Illumina

Nukleinsäure-Messgeräte

- NanoDrop® 1000 Spectrophotometer, Peqlab Biotechnology GmbH
- NanoDrop® 8000 Spectrophotometer, Peqlab Biotechnology GmbH
- TapeStation2200, Agilent GmbH

Pipettierhilfen

- accu-jet® pro, BRAND GmbH & Co. KG
- Eppendorf Research® 0,1-2,5µl; 0,5-10µl; 2-20µl; 10-100µl; 20-200µl; 100-1000µl; 500-5000µl, Eppendorf GmbH
- eLINE Single Electronic Pipette 5-120µl, Sartorius AG
- Eppendorf Research pro 20-300µl
- Finnpipette® 16-Kanal, VWR International GmbH
- Multipette® plus, Eppendorf AG
- Transferpette® 0,5-10µl; 2-20µl; 20-200µl, BRAND GmbH & Co. KG
- Transferpette® S-8 0,5-10µl; 2-20µl; 20-200µl, BRAND GmbH & Co. KG
- Transferpette® S-8 0,5-10µl; 2-20µl; 20-200µl, Eppendorf AG

Pipettierroboter

- Biomek® NX MC Laboratory Automation Workstation, Beckman Coulter GmbH
- Biomek® NX S8 Laboratory Automation Workstation, Beckman Coulter GmbH

pH-Wert Messgerät

- inoLab® pH 720, Wissenschaftlich-Technische Werkstätten GmbH

Reinstwassersystem

- Milli-Q A10 Synthesis, Merck KGaA

Schüttel-, Misch- und Heizgeräte

- Überkopfschüttler REAX 2, Heidolph Instruments GmbH & Co. KG
- Thermomixer comfort, Eppendorf AG
- Thermomixer compact, Eppendorf AG
- Titramax 101 kreisförmig vibrierender Plattformenschüttler, Heidolph Instruments GmbH & Co. KG
- Unimax 1010 rotierender Plattformenschüttler, Heidolph Instruments GmbH & Co. KG
- Vortex Genie 2, Scientific Industries Inc.
- VWR Signature™ High-Speed Microplate Shaker, VWR International GmbH

Spannungsgeräte

- Powerpac Power Supply, BioRad Laboratories GmbH

Thermocycler

- C1000 Thermal Cycler, BioRad Laboratories GmbH
- Mastercycler® pro, Eppendorf GmbH
- PTC-100, MJ Research
- PTC-200, MJ Research

Trockenschrank und Inkubatoren

- Brutschrank, Memmert GmbH
- T 20 P, Heraeus GmbH

Waagen

- TE3102-S, Sartorius AG
- TE3132-DS, Sartorius AG

Zentrifugen

- Biofuge fresco, Heraeus GmbH
- Biofuge picco, Heraeus GmbH
- Biofuge stratos, Heraeus GmbH
- Concentrator plus, Eppendorf AG
- Megafuge 1.0, Heraeus GmbH
- Megafuge 1.0 R, Heraeus GmbH
- neoLab-Mini-Zentrifuge Spectrafuge, neoLab
- Vakuumzentrifuge Speed Vac Plus SC110A, Thermo Electro Corporation

Chemikalien und Enzyme

Soweit möglich, wurden Chemikalien mit der Qualitätsbezeichnung „pro analysis“ verwendet.

- Agarose low EEO (Agarose Standard), AppliChem GmbH
- Alconox, Alconox Inc.
- Bromphenolblau, Sigma-Aldrich
- Dimethylsulfoxid (DMSO), Sigma-Aldrich
- DNA ladder 1kb und 100bp, GeneCraft
- dNTPs [10mM], Labomedic GmbH
- Essigsäure (100%), Merck KGaA
- Ethylendiamintetraessigsäure (EDTA, reinst), Merck KGaA
- Ethanol absolut (EtOH, 100%), AppliChem GmbH
- Ethanol (EtOH, 96%), WALTER CMP GmbH & Co. KG
- Ethidiumbromid (EtBr-Lsg, 1%), Merck KGaA
- FicollTM PM400, Amersham Biosciences GmbH
- HotStar Taq DNA-Polymerase [5 U/μl], Qiagen GmbH
- HPLC Wasser, Merck KGaA
- Natriumacetat (NaAc), Merck KGaA
- Natriumhypochlorit (12%), Carl Roth GmbH & Co. KG
- Natronlauge (NaOH-Lsg, 0,1M), AppliChem GmbH
- QIAzol Lysis Reagent, Qiagen GmbH
- 10 x TBE Puffer, Life Technologies GmbH
- Tris (≥99,9%), Carl Roth GmbH & Co. KG
- VWR-Taq DNA-Polymerase [5U/μl], VWR International GmbH

Lösungen

Wenn möglich, wurden alle Lösungen mit demineralisiertem, sterilfiltriertem und autoklaviertem Wasser angesetzt.

Bromphenol-Auftragspuffer

- 10ml 10 x TBE, 10ml 0,1% Bromphenolblau, 40ml 20% Ficoll (8g) ad 100ml dest. Wasser

Ethanol 70%

- 729ml EtOH (96%), ad 1000ml dest. Wasser

Ethanol 85%

- 885ml EtOH (96%), ad 1000ml dest. Wasser

Natriumacetat (NaAc) 3M, pH 5,2

- NaAc mit Essigsäure (100%) auf pH-Wert von 5,2 eingestellt

Tris-EDTA (TE⁻⁴)

- 0,1mM EDTA, 10mM Tris-HCl, pH 8,0

Kommerzielle Kits

- Agencourt AMPure XP Kit, Beckman Coulter
- Agencourt CleanSEQ Kit, Beckman Coulter
- BigDye® Terminator Cycle Sequencing Kit 3.1, Life Technologies GmbH
- Chemagic DNA Blood Kit special, PerkinElmer Chemagen Technologies GmbH
- EDTA-haltige Monovetten von 1,2ml bis 10ml, SARTEDT AG & Co.
- HiSeq Rapid SBS Kit v2, Illumina
- HumanOmniExpressExome 8 v1.2 und v1.3 BeadArray, Illumina
- iPLEX™ Gold Reagent Kit, Sequenom® GmbH
- LightCycler® 480 SYBR Green I Master, Roche
- Oragene® DNA (OG-250) und (OG-500) DNA Genotek, OraSure Technologies Inc.
- PowerPlex® 16 HS System, Promega GmbH
- QIAamp DNA Blood Maxi Kit, Qiagen GmbH
- QuantSeq 3' mRNA-Seq Library Prep Kit, Lexogen GmbH
- REPLI®-Midi Kit, Qiagen GmbH
- RNeasy Plus Mini Kit, Qiagen GmbH
- SeqCap EZ Human Exome Kit v2.0, Roche
- SpectroCHIP® Arrays & Clean Resin Kit, Sequenom® GmbH

Software und Datenbanken

- Assay Design 3.1, Sequenom® GmbH
- BeadScan, Illumina

- BiomaRt R Package, Bioconductor
- Biomek® Software 3.2, Beckman-Coulter GmbH
- Cartagenia Bench Lab CNV, Cartagenia GmbH
- 2D CYPHER™ Pilot Databases, Thermo Fisher Scientific GmbH
- 3130xl DataCollection v3.0, Life Technologies GmbH
- dbNSFP (<https://sites.google.com/site/jpopgen/dbNSFP>)
- dbSNP database (<http://www.ncbi.nlm.nih.gov/projects/SNP/>)
- DECIPHER (<https://decipher.sanger.ac.uk>)
- DGV (<http://dgv.tcag.ca/dgv/app/home>)
- edgeR v3.16.5, Bioconductor
- ENSEMBL genome browser (<http://www.ensembl.org/index.html>)
- Exome Variant Server (EVS) (<http://evs.gs.washington.edu/EVS/>)
- FastQC v0.11.4 (<http://www.bioinformatics.babraham.ac.uk/projects/>)
- FluidX IntelliCode™ Decoding Software, FluidX GmbH
- GeneCards (<http://genecards.org>)
- 1000Genomes Browser (<http://browser.1000genomes.org/index.html>)
- GenomeStudio v2.0.2, Illumina
- Haploview Version 4.2, Broad Institute
- HapMap genome browser (<http://hapmap.ncbi.nlm.nih.gov>)
- KEGG (<http://kegg.jp>)
- limma v3.30.13, Bioconductor
- MySequenom, Assay Design Suite (<https://mysequenom.com/Home>)
- MGI (<http://www.informatics.jax.org>)
- NanoDrop® ND-100 v3.3.0, Peqlab Biotechnology GmbH
- NanoDrop® ND-8000 v2.2.1, Peqlab Biotechnology GmbH
- NHLBI Exome Variant Server (<http://evs.gs.washington.edu/EVS/>)
- NCBI (<http://www.ncbi.nlm.nih.gov>)
- Primer3 v0.4.0 (<http://www.frodo.wi.mit.edu>)
- OMIM (<https://www.omim.org>)
- Pubmed (<http://www.ncbi.nlm.nih.gov/pubmed/>)
- Protter (<http://wlab.ethz.ch/protter/start/>)

- QuantiSNP (<https://sites.google.com/site/quantisnp/>)
- R, R-Studio
- RegulomeDB (<http://regulomedb.org>)
- RT-Workstation 3.3, Sequenom® GmbH
- SAS Software version 9.1
- SeqMan II Version 5.0, DNAST AR Inc.
- SmartScanTM 96, Thermo Fisher Scientific GmbH
- SpectroAQUIRE, Version 3.3.1.2, Sequenom® GmbH
- SpectroPOINT, Sequenom® GmbH
- STRING (<https://string-db.org>)
- Subread v1.5.1 (<http://subread.sourceforge.net/>)
- Typer 3.4/ 4.0, Sequenom® GmbH
- UCSC Genome Browser (<http://genome.ucsc.edu/>)
- UniProt (<http://www.uniprot.org>)
- Varbank (<https://varbank.ccg.uni-koeln.de>)
- ZFIN (<http://zfin.org>)

Appendix A2: Auflistung der im Rahmen dieser Arbeit verwendeten Primersequenzen.

Dargestellt sind die 5'-3' Sequenzen der Primersequenzen (F: Forward; R: Reverse). Das Design der Primer wurde nach hg19 durchgeführt.

Primer für Replikationsstudie bei Patienten mit BEEK

SNP-ID	F	R	UEP
rs72748303	ACGTTGGATGTGGAAGCGGGACATTCAATC	ACGTTGGATGGGTGAACAGGAAAAACGGG	GGACTCTTGCTTTGCC
Gyc-Marker	ACGTTGGATGAGTCAGCTCAGGAGGAAATC	ACGTTGGATGGGAGTGCATCCTTTAGGAAC	CCCTCCTCGGGAGTTCT
rs56189125	ACGTTGGATGTGGAATGCTGCCAAGAAAGG	ACGTTGGATGTGGCCATCTTGAGAGGTTG	CCATTTCTTCCCCACTA
rs6874700	ACGTTGGATGCTTCACTGTTGTTATTGGAC	ACGTTGGATGCAAGCCTGGTAACTAGACTC	GTGTGGTGATTCTATAGAGA
rs1514921	ACGTTGGATGAAGGCACAGAGAGGCAGTAG	ACGTTGGATGAACTGGCCCTTAGGGTACAC	TGTTATCCCATTCAACTAATAG
rs16917077	ACGTTGGATGAGGTTAACTAGCACCCCTCAG	ACGTTGGATGAGTGATCTTGCCGAGGGATG	GGGACCTCAGGATGCATAGTAC
rs10119066	ACGTTGGATGAAATCACACTGCGAGCACTG	ACGTTGGATGATATTCCAGGCTCACTGAAG	GGCACTGAAGAAAAATCTCCAGA
rs57086087	ACGTTGGATGGGTTCTACCAAGTTTTCAAC	ACGTTGGATGGCTTAATTCCAAGCCAAGTG	TACCTAGAAATCTACTTCAGTTTA
rs1475601	ACGTTGGATGCCTGTCTAAACACATAAAG	ACGTTGGATGTGTACGATCAATTAACCAGG	CCCCCTTTTTTCATTTTGTGAGCAT
rs7689350	ACGTTGGATGCCCAAAGATCTTTATGGAGTG	ACGTTGGATGCTAGTATTTGCTTACTGGT	GGTTAGTTTTCTTTGTTCTCAAGAT
rs117622209	ACGTTGGATGTGAGGAAGACTGTGCAGAAC	ACGTTGGATGTCTGCTGCCTCAACACAAC	GAGGGTGTGAATAAGCTTGTATCA
rs1514351	ACGTTGGATGCACTATTGTAGAAGGAAACG	ACGTTGGATGGTATAGCAGTATATTTGCGG	TATTTTAGAAGGAAACGAAAGGAGTTA
rs79145914	ACGTTGGATGCCAGCACACCGATTTAAAC	ACGTTGGATGTATGGGTCCCAAAGTGTTG	GTACACCTGGGGACTTTCTTTAGTTAA

Primer für Kandidatengeneanalysen bei Patienten mit BEEK***ISL1***

Primername	Sequenzen	Temperatur
ISL1-i1F	GGGCTCGGTGTGCTGTTCT	63,8 °C
ISL1-i1R	ATTGCCGCAACCAACACATAG	63,1 °C
ISL1-i2F	TCATAATTCCGGGGTATTGAGG	62,0 °C
ISL1-i2R	CTTGTTAATAGCCACCGGACATT	61,4 °C

EFNB1

Primername	Sequenzen	Temperatur
EFNB1-1F	CAGCGCGGCAGAGGAAGG	67,8 °C
EFNB1-1R	CCACATGCACTCCGGGGC	68,5 °C
EFNB1-2F	ACCCAGTAGGCCAGCCC	64,7 °C
EFNB1-2R	AGGGGATCAGATCCAGCAGG	63,8 °C
EFNB1-3F	GCAGTTTCTCTCCCTGTTGGC	63,5 °C
EFNB1-4R	GCTTGCATTTCTTCAGGGACAG	63,4 °C
EFNB1-5F	AAGAAATGCAAGCTGGGCCTG	65,8 °C
EFNB1-5R	CACAGCTGGCAAGGGGAGG	66,5 °C

SLC20A1

Primername	Sequenzen	Temperatur
SLC20A1-2F	GCTCGGTTTCTGTGCCGTAG	63,2 °C
SLC20A1-2R	TTATGCACAAAAATGGAACCACC	62,9 °C
SLC20A1-3F	GGGTGAGCAAATTTGTATCATGG	62,6 °C
SLC20A1-4R	GGCCGCATACCTATACACCAAT	62,1 °C
SLC20A1-5F	ATTGGCACTATAATTCCCAAACC	60,3 °C
SLC20A1-5R	AGCGTACACTAGCCTGCAGC	60,8 °C
SLC20A1-6F	GTCCCCAGAGGAGGGTTTAC	59,8 °C
SLC20A1-6R	TAAAAACAGTTTGGACAACAAACC	59,4 °C
SLC20A1-7F	GGTGGAGGAAAGGAGACGTG	62,0 °C
SLC20A1-7R	AAGCCTCCTCACCCTGCAG	62,9 °C
SLC20A1-8F	TACCTGCAGTGGTGAGGAGG	61,3 °C
SLC20A1-8R	CACTACCACAAGCACAGGCC	61,7 °C
SLC20A1-9F	TTCTGTTGTGTATGTTGGGGTTC	61,0 °C

SLC20A1-9R	GTCTGACCAAAATGTTTAAGGGC	61,1 °C
SLC20A1-10F	AGGAAACTGGGCATCTGTGG	62,4 °C
SLC20A1-10R	ACCATTTGAAAGACACACATTGG	61,0 °C
SLC20A1-11F	TCTTGTCCAGGGGTCTGGG	63,5 °C
SLC20A1-11R	GGGGACCTCAGCAGGAGC	63,4 °C

Primer für qPCR bei CNV-Analyse in Patienten mit VATER/VACTERL

Primernamen	Sequence	Temp
SMYD3-in6-ex6-F	ACAAAGAGGCCAGCAACAG	62,7 °C
SMYD3-in6-ex6-R	TCAAGGCATGTTTCCCAAG	63,2 °C
SMYD3-ex5-F	AGGCTTCCTCCAACAGGCC	63,9 °C
SMYD3-ex5-R	GCAATTCTGAACGGCTTCCC	64,1 °C
SMYD3-in5-F	GCTGGGTTTCTCTCTACC	60,1 °C
SMYD3-in5-R	AGCTCGTTTTCTCTCACCTGG	60,9 °C
FAM170A-ex1-F	GGGAGGTGGACATTAAGCATCT	61,6 °C
FAM170A-ex1-R	CGCACATACCTCCTCCCTTC	61,9 °C
FAM170A-in3-F	AGGAAGTGAGCAAAGCCTGG	61,8 °C
FAM170A-in3-R	CCCTGAGCCAATAAAGCCC	61,8 °C
FAM170A-3'-F	ACAACTTGTTATCATACGGCCC	60,1 °C
FAM170A-3'-R	TCCCTCTACATACGGGGAGTC	60,3 °C
CNTNAP2-in13-F	CCTACCTCCACATCAATTCGG	61,6 °C
CNTNAP2-in13-R	CGGCAACTCTTCTCTACCCATAG	61,5 °C
CNTNAP2-in14-neu-F	TCAGAGAGCTTCACTGTGTTGAG	59,8 °C
CNTNAP2-in14-neu-R	CAGGGTCTAGGTATGAGGAAAGG	60,4 °C
CNTNAP2-in15-neu-F	GTTTCGCTGTTTACCTCGTCC	59,7 °C
CNTNAP2-in15-neu-R	TCCTGGAATGTTTAGGTGGG	59,8 °C
FBP2-ex7-neu-F	ACAGAAGAGGGGCATGTGGG	61,1 °C
FBP2-ex7-neu-R	GGCAATTCACCAGCGAGTC	61,8 °C
FBP2-in1-neu-F	AGAGCCTGGGACCCTAAAAC	59,6 °C
FBP2-in1-neu-R	TAAGGGAGGCAAGGCTGAG	60,5 °C
FBP1-in2-neu-F	CATGCCAGACCTACGAGAG	60,8 °C
FBP1-in2-neu-R	TCTGTGTTGGTTCTCTGTGTG	61,2 °C
ANKS1B-ex1-neu-F	GACTCTCTCCTCTTCGGG	61,8 °C
ANKS1B-ex1-neu-R	ATTCAGGGGAGTGGCTGTG	61,1 °C

ANKS1B-3'-F	AGAGACATCATATGCATACCCTGG	61,4 °C
ANKS1B-3'-R	AGAGAGGTTTGGTGCCCTTG	61,6 °C
UHRF1BP1L-ex20-F	AGGTTTCCCATTACCTGTTCCC	62,4 °C
UHRF1BP1L-ex20-R	GATTCAGTGCTGCTGACCAGTG	63,3 °C
GPC5-in7-L-F	GTGCATTTCTACAGCACACCTG	60,8 °C
GPC5-in7-L-R	CCAAGGGACCTTCTGTACCAGT	61,7 °C
GPC5-in7-M-F	AGTGGGTGCATTCTACTTCCAGAT	62,4 °C
GPC5-in7-M-R	CCCTCTGGCAATTTTCATCATAAC	62,9 °C
GPC5-ex8-F	TGACCCACTGTAAGTTAGCCACT	60,1 °C
GPC5-ex8-R	AGAGAGAGATGGGAGCAGAGG	60,1 °C
NALCN-AS1-in2-F	CATACTAAGACGGGCGGGAG	61,9 °C
NALCN-AS1-in2-R	GTGCTGCTCTCAAATCTGCAC	61,1 °C
NALCN-ex44-F	AACTTCAGAACCGCAGGAAGC	63 °C
NALCN-ex44-R	ACAGCCCCAAAACCAATAAGC	62,4 °C
NALCN-ex31-in30-F	GGCTCTTCAGGTCTTCCCATC	62,4 °C
NALCN-ex31-in30-R	CCATCTGCCATGTTCTTGGTG	63,7 °C
VPS13C-in42-F	CAATCAAAGTACGCGCCCTAAATC	61,2 °C
VPS13C-in42-R	GACAGATACTTTCCGTGCTTTTG	60,2 °C
VPS13C-in32-F	GACTGACACTTCTCCTAGCCCAT	61 °C
VPS13C-in32-R	AGAATCTGACCATCCCAATTGAC	61,4 °C
VPS13C-in16-F	GAAGCCACCAGAACTGCATG	61,8 °C
VPS13C-in16-R	CTGTCTAGTTAGCAAGCCAGC	60,6 °C
ZSCAN2-in2-F	GGTGTTCGGAGGAGGAGAAG	62,4 °C
ZSCAN2-in2-R	TGGAGCTGACTTTGCACACAGT	63,3 °C
ZNF592-ex4-neu-F	GTTTCCGGTTCCAGAGCTG	60,8 °C
ZNF592-ex4-neu-R	AGGTTCCACTGTGTTCTGCC	60,2 °C
PDE8A-in6-F	GCAGCATGGCAAGTGAGAAC	62 °C
PDE8A-in6-R	GTACGAGGGGAGTTGCCACT	61,5 °C
CDH13-in7-L-F	CTCCACTGAGAACTAATTGCCTG	60,3 °C
CDH13-in7-L-R	TATCCATTTTGAGCTCCACTGAG	60,6 °C
CDH13-in7-M-neu-F	ATCGGGGTGTAGCAATCTTG	60 °C
CDH13-in7-M-neu-R	GATAAGTGGGGTTGGCACAG	60,4 °C
CDH13-in7-R-F	ACCTTACATCCACGGTGAGC	60 °C
CDH13-in7-R-R	CCACATTTTGCTGGGAGAC	59,1 °C
MOY19-in23-ex23-F	TCATCCCCAAACCAGCTCAC	63,4 °C

MOY19-in23-ex23-R	TTCTCCACACTCTGCCGGTC	63,8 °C
ACACA-in43-ex43-F	CCAATGGGATTTGTGAGTGTG	61,2 °C
ACACA-in43-ex43-R	ACTGACTCCAGGACAGCACAG	60,5 °C
DDX52-in10-F	CGTCACCATTCTCTCCAATTG	60,5 °C
DDX52-in10-R	GTTGAGGCATCTGTTACCATAGTG	60 °C
DCX-ex3-neu-F	AAGTCCTTGTTCTCCCTGGC	60,6 °C
DCX-ex3-neu-R	CCAAGAATGTCAATCCCAACTG	61,6 °C
TRPC5-ex4-F	CAAGGCAATATCATCCATTCCAC	62,6 °C
TRPC5-ex4-R	GCCACACAGCATCCTATTTGAC	61,8 °C
AMOT-in1-neu-F	AAGTAGGACGGGAAAGCAGC	60,8 °C
AMOT-in1-neu-R	GAGTGAATCGATCTAGGCGC	59,9 °C
USP37-ex20-F	TGACTCAGAGATGGTGAAGCATC	61,7 °C
USP37-ex20-R	TGCCCAATAGAGCCTGACAAG	62,6 °C
USP37-in11-F	CCCTAACTCTCAGCACCCAC	59,7 °C
USP37-in11-R	GCTAGGAGACAATACCACCACC	59,9 °C
USP37-in10-F	TCGGTTTCCTTGCTATATTCCAC	61,5 °C
USP37-in10-R	GCTTTACTGTTACCTGCATGGG	60,9 °C
PARM1-5'-1-F	CCTTCTTGTGTGGTCATGTGC	61,6 °C
PARM1-5'-1-R	CTGTAACCACCTGCTGGCAG	61,8 °C
PARM1-5'-2-F	AAAGTGCAGTTTAGCCCCATGG	64,5 °C
PARM1-5'-2-R	CAAACACTTGGGCATGCTGG	64,8 °C
PARM1-in1-F	GACTTCTTGTTGCCGTTGGG	62,9 °C
PARM1-in1-R	ACCGCCAAGACTCGAGAAAG	61,8 °C
CDH12-in3-L-F	TTTCCAGAAGAGAGGCAGGTAG	60 °C
CDH12-in3-L-R	ACTTAATTTACGGAAAGTTGCC	59,8 °C
CDH12-in3-M-F	TCTAGCCAGAGAACAGCAGAGG	61,2 °C
CDH12-in3-M-R	TCCTAAGCATTCTAACCTCCTGC	61,1 °C
CDH12-in3-R-F	TGGCACCTAAAGTAGATGGGC	61,4 °C
CDH12-in3-R-R	TGAAACCCGACTCTGCTGTC	61,4 °C
NREP-in2-L-F	TCCTCCATGGGTCTTCTGGTT	63,1 °C
NREP-in2-L-R	TGGTTGTGAAAGTAATGCCCTG	62,1 °C
NREP-in2-neu-F	GAGACTCCTAGACTTCTTGCCC	60 °C
NREP-in2-neu-R	GAAACTGAGGTGGCAGCAAC	60,8 °C
NREP-AS1-in3-F	GCTTCACTTACCTGGGATGTG	60,9 °C
NREP-AS1-in3-R	CTGTTTCAGCCTCAGTGTGACT	60,9 °C

ZNF713-in1-F	ACTCTGTTGCTGAAAACCTTGCG	61,9 °C
ZNF713-in1-R	GTGCATGGGTGTAAGACAAAGC	61,8 °C
GBAS-in1-ex2-F	AACGGTGCCTAGAATAGCTGCT	61,6 °C
GBAS-in1-ex2-R	TGTCTTCTCGAGATCTGTTGCTG	62 °C
PSPH-in1-F	CCTATCTTTGGTCTCTCTCCGC	61,6 °C
PSPH-in1-R	TGGCATTTTAACTCTCCCGG	62,2 °C
SORCS1-in20-F	TGAGTCCCTCTATGCCTCCC	61,5 °C
SORCS1-in20-R	CATCTCCCCAGCCCTGAAT	62,3 °C
SORCS1-in3-F	GTAATGGGGTGAAGGCAATGC	63,7 °C
SORCS1-in3-R	TTATGGTTGGTGGGCAGTGG	64 °C
SORCS1-ex2-in1-F	GCGAGCTCTCTGTGATGCTC	61,4 °C
SORCS1-ex2-in1-R	AATCCTAAGCCCTCCCTTGG	61,6 °C
SGK223-in3-F	GTGGTCAGTCCTGTTTTGGCT	61,5 °C
SGK223-in3-R	GGGAGACTTAGTTGGGACGC	61 °C
SGK223-in2-F	CTGTCCTCATCAGTGAGGCCT	61,8 °C
SGK223-in2-R	CTGAAGCTGGCTCTTCACACC	62,4 °C
SGK223-ex2-F	ATCCCGGTCCCAACTTTG	60,7 °C
SGK223-ex2-R	GAGTCACAGCTGGGGAAGAG	60 °C
LOC389043-ex2-F	GACCATTCTGTGTGCACC	60 °C
LOC389043-ex2-R	GGACCCTGAAGAATCTGGAATC	61,2 °C
MZT2A-in2-F	AAAACAAGGAAGAGCACCAACC	61,7 °C
MZT2A-in2-R	CACACACTCACCTCTCTCCAG	61,5 °C
LOC150776-in3-F	TGACTTACTTAGCTGGCCCAAG	60,8 °C
LOC150776-in3-R	CTGAGACCTACGAGATCAGCC	59,1 °C
vor-XXYLT1-5UTR-F	AGTGACCTTCGGATGAGATCTTAG	60,1 °C
vor-XXYLT1-5UTR-R	TGCCTTTCAGATGCCTTGAC	60,3 °C
XXYLT1-ex5-F	GTGGTAGATCTTGACGTGGCC	61,9 °C
XXYLT1-ex5-R	CTGGACTGTACCTGGAACCG	60,6 °C
XXYLT1-in4-F	CAACAGAGTCCATCCCTGAATG	61,8 °C
XXYLT1-in4-R	CGCTGATAGTTGAAACAGGTCTC	60,3 °C
AGMO-in5-F	TTCTGATCCTAGGCCCTGG	60,2 °C
AGMO-in5-R	AGGAGCAAAAGATTTTCACCG	60,6 °C
AGMO-in4-F	ACTTGGGGAATGTCGTGTTG	60,8 °C
AGMO-in4-R	TCAACCAAGACACTTCTGATTC	60,5 °C
AGMO-in3-F	ACTAGAGCCCTAATCGGTCATG	59,6 °C

AGMO-in3-R	GTCAGAGCTGTATCACCTGCAG	60,1 °C
FXYD4-in5-F	TAAGAGCTGATATCCCACCCC	61,5 °C
FXYD4-in5-R	CCCATCCCCTTTGTCTACACTAG	61,1 °C
HNRNPF-ex4-F	GGACATGAACTTCAGAGGGG	59,5 °C
HNRNPF-ex4-R	AGAAGGCTCTAGGGAAACACAAG	60,3 °C
ZNF487P-in1-F	CACTTGTAGTCAGGGTTCGTGG	61,9 °C
ZNF487P-in1-R	CAAGTCTGTAACTCCCTGTGCC	62,3 °C
EPSTI1-in5-F	CAAAATCCTAAGGTTCAACTGGC	61,2 °C
EPSTI1-in5-R	TTTCAGTGCAGTGCCCATG	61,9 °C
DNAJC15-ex1-F	TGATTCAAGAACGTCCCAGTG	61,1 °C
DNAJC15-ex1-R	TGCTATTAAAGGCGAGGTTCC	60,6 °C
DNAJC15-in5-F	TTGGGGAAGTATGAGAATGGG	61 °C
DNAJC15-in5-R	TTCATGAATTCGGAAGCACTG	61,2 °C
ABR-in2-F	TCCCTTTTACCAAGTCCTTCG	60,5 °C
ABR-in2-R	GACAAGATCAGGAGCCAAAGTG	61,2 °C
ABR-in1-F	AGAACAAGTCTTCCATCTACTCG	60,2 °C
ABR-in1-R	TTGGCTACTGTTGTTCTTCGG	60,3 °C
TUSC5-in1-F	TGAGGACTGCTCAAGGGTGTAC	61,6 °C
TUSC5-in1-R	TTATGGACTGGGAGGATAAGGC	61,5 °C
TBCD-in38-F	TAAGAGCAGTGCAGTGAACCTC	59,7 °C
TBCD-in38-R	GAAACCATTTGTGCCGTTTC	60,4 °C
B3GNTL1-in5-neu-F	CAAATGGCCCATATGCAAG	59,9 °C
B3GNTL1-in5-neu-R	GAATGGAGCAGGTTCAAGACG	60,8 °C
METRNL-in1-F	CTGTTGCTGGGTGTTTCTTCTAAG	61,5 °C
METRNL-in1-R	ACAGCAAGTGGGAAAGTTCATG	61,4 °C

Primer für WES in Patienten mit EA/TEF

Primernamen	Sequenzen	Temperatur
PPIP5K2-7F	CAGCAACCTGGATGAATCTTAAG	60,1 °C
PPIP5K2-7R	GAGTCCATAGGCTTCAGACTGC	60,4 °C
SLC5A2-6F	GGAAGGGGAACTCTTCAAATTC	61,4 °C
SLC5A2-7R	TCTGTCTGGTTTAGCACCAACC	61,4 °C
EEF1D-1F	TTGGGGTCTCTGCACCCATG	66,6 °C
EEF1D-1R	CCAGCCTGGGATGGCTGTC	66,5 °C

HPS3-6F	GAAGATACTGACCCAGATCCATG	59,8 °C
HPS3-6R	ACTGTCAACTAAGCCATGAATCAC	59,6 °C
DNAH6-54F	GGTTCTGAAATGGGAATTAATGC	60,9 °C
DNAH6-54R	GAAGTCAGAAAAATGAACCACCC	61,1 °C
TANC2-12F	TGAAGTTCCCAACCCAGTCTTC	62,6 °C
TANC2-12R	GCCTTGCCTTGAGAAGAAACAC	62,5 °C
UBA3-14F	TGGTCTATTATGAAAAGCAGGCTT	60,5 °C
UBA3-15R	CTAGGGTGGCTGTGATGGCT	62,5 °C
NFX1-8F	TTTGGCCTCCCAGATTACAGG	63,4 °C
NFX1-9R	GGGCAAATGTATGGGTTTTTCAC	62,5 °C
GGT6-3F	AGACCTGGGGGTGGAGGTG	64,7 °C
GGT6-3R	GTGTCTGCTCTGCTCCTGCC	63,7 °C
CLP1-3F	AGAGGAACAAAGTAGCCATGGG	61,7 °C
CLP1-3R	TCTCAGCCCATCTGCCTCTAG	62,0 °C
TPP2-13F	TAGAATTGTCCTCTTAGTTGGGG	58,7 °C
TPP2-13R	GTATTCTCTTCACAACACTGGCC	60,1 °C
NPR2-3F	TTTTACCAAGTTTCCCTGTCCTG	61,5 °C
NPR2-3R	TATTCCTAAAGTGCCATCCCTTC	60,7 °C
PIGC-2F	AACACAGCATGGAATTTCCAGG	63,3 °C
PIGC-2R	AGGATAAGCACCCCTCACCACC	62,5 °C
STAB1-56F	CCTGGGGAGGAGTGCTGGG	67,6 °C
STAB1-57R	TCCACCCACCCTTGGTGCC	68,8 °C
ZFHX3-1F	GGACAGGCCCCATGAGGAG	65,3 °C
ZFHX3-1R	TCATTTGGTTCTGGTGCTGTGG	65,4 °C
ADGRD1-10F	TTGTGTGCCTGGTGACCTG	66,8 °C
ADGRD1-10R	TCATGGCCGAACACACCTGC	67,8 °C
PCDHA10-1F	GCCCAGGTTCTCCGTAACAG	62,0 °C
PCDHA10-1R	TCAGCAGAGAAACAGATCCGG	62,8 °C
KIAA0556-19F	GCCGAATAGCATTCTCTTCC	62,2 °C
KIAA0556-19R	CATTACCTCACAGGGTGGCC	62,6 °C
APOL2-5F	TTCCATATGGAATAAGTACTGAGG	57,8 °C
APOL2-5R	CCAGAGGTAGTGCCAACAGAG	59,9 °C
CELSR1-3F	CTGTGTGGCTGGGGTGTGTC	65,1 °C
CELSR1-3R	CTAGAGGGATGGGGTGCACG	65,1 °C
SLC6A8-5F	AGCCTTGCCCCTAGCTTCCT	63,7 °C

SLC6A8-5R	CCTGCTGCAGCCTCTCCAC	64,7 °C
FANCB-3F	GGTTTTATTGGGACTAAAGGAATG	59,2 °C
FANCB-3R	AAAGGTCTGAGACCACCAACTG	60,6 °C
MTA3-6F	GGTGTCTTACCTTGGAGTTATGG	58,9 °C
MTA3-6R	CAGTAAAATGGCAAGTGCTTATAAC	58,5 °C
CHD7-18F	CCATCACATTGGAATGAGGGTT	63,1 °C
CHD7-18R	ATAAATGGTTTGGGCAAGGGTG	64 °C
TRPS1-4F	TGCTAGAACATTATGGCAAGCAG	61,6 °C
TRPS1-4R	TAATTTCTCCAAGGTGCTTTTCTG	60,9 °C
PLEC-31F	GCAGAGGAGCAGGCACAGGC	68,1 °C
PLEC-31R	GCTGCTGTGCCAGGTCCTCC	67,9 °C

Primer für Kandidatengeneanalyse EEF1D bei Patienten mit EA/TEF

Primername	Sequenzen	Temperatur
EEF1D-ex3-F	CCCACATGGGCTTGACTTC	61,5 °C
EEF1D-ex3-R	CTCCTGCATTCAACTCTGCTG	61,1 °C
EEF1D-ex4-5-F	CTTTGTGGAGAACCGCTGG	61,8 °C
EEF1D-ex4-5-R	GCAGCATCAGGAAAAATCAGAC	61,1 °C
EEF1D-ex6-7-F	GAGCTGTAGTTTTCTTCCGTGAG	62,7 °C
EEF1D-ex6-7-R	CAGCTCATCACTGCTGCTGG	63,3 °C
EEF1D-ex8-F	GCCCTTCCTCCAGTCAAAG	59,8 °C
EEF1D-ex8-R	GCCTCCAAGTTTGTGTTTATCC	59,9 °C
EEF1D-ex9-10-F	GACACACAGTGGTTGAGGGTAG	59,5 °C
EEF1D-ex9-10-R	GTACGAGCAGGAAATGACCAG	59,7 °C

Appendix B: Ergebnisse

Hardy-Weinberg-Gleichgewicht-Test für Fall-Kontroll-Studie in verschiedenen Kohorten

SNP	ESP		GB		AUS/DE		Schweden		Italien	
	P-Fall	P-Kontroll								
rs1475601	8.62E-01	9.64E-01	8.69E-01	8.17E-01	9.33E-01	7.35E-01	5.92E-01	6.41E-01	8.69E-01	7.33E-01
rs72748303	6.49E-01	3.80E-01	8.71E-01	9.09E-01	8.00E-01	7.71E-01	5.51E-01	6.98E-01	6.03E-01	6.01E-01
rs79145914	5.79E-01	7.76E-01	9.36E-01	8.17E-01	7.97E-01	7.39E-01	7.75E-01	6.98E-01	8.03E-01	5.59E-01
rs7689350	3.65E-01	9.97E-01	6.69E-01	3.51E-01	5.79E-01	9.77E-02	8.74E-01	7.71E-01	4.19E-01	3.31E-01
rs6874700	2.40E-02	1.31E-02	2.06E-01	2.87E-01	3.71E-02	3.43E-01	8.62E-01	8.22E-01	8.85E-01	1.45E-01
rs1514351	6.15E-01	2.79E-02	5.81E-01	1.80E-01	1.36E-01	4.11E-01	9.25E-01	6.33E-01	8.85E-01	6.39E-01
rs117622209	7.20E-01	8.84E-01	8.71E-01	7.71E-01	8.68E-01	6.12E-01	8.10E-01	8.17E-01	2.50E-02	5.57E-01
rs56189125	1.00E+00	5.55E-01	9.36E-01	9.55E-01	1.00E+00	7.71E-01	1.31E-01	9.70E-01	9.35E-01	9.11E-01
rs57086087	3.24E-01	7.23E-01	8.05E-01	2.53E-01	4.06E-04	6.49E-01	9.12E-02	8.42E-01	6.36E-01	8.28E-01
rs10119066	5.11E-01	3.40E-01	6.73E-01	5.48E-01	7.32E-01	4.11E-01	2.64E-01	2.92E-01	8.03E-01	5.59E-01
rs16917077	7.20E-01	8.84E-01	7.39E-01	7.71E-01	3.60E-03	7.07E-01	1.62E-02	2.26E-07	8.69E-01	6.45E-01
rs1514921	4.45E-01	7.93E-01	9.01E-01	1.17E-01	5.85E-01	2.13E-01	3.13E-01	2.75E-01	5.51E-01	5.85E-01

ESP: Spanische Kohorte; GB: Britische Kohorte; AUS/DE: Australische/Deutsche Kohorte; Schweden: Schwedische Kohorte; Italien: Italienische Kohorte.

Danksagung

Mein besonderer Dank gilt Herrn PD. Dr. Heiko M. Reutter, der mir dieses äußerst interessante Thema überlassen und meine Arbeit stets durch Ratschläge, Hilfestellungen und fachliche Diskussionen unterstützt hat. Bedanken möchte ich mich an dieser Stelle auch für das in mich gesetzte Vertrauen, neben der Wissenschaft, den Laborablauf koordinieren und die Studenten betreuen zu dürfen. Zudem möchte ich mich für die zur Verfügung gestellten Labormittel, den Arbeitsplatz und letztendlich für die Begutachtung dieser Dissertation bedanken.

Herrn Prof. Dr. Michael Pankratz gilt ebenfalls mein besonderer Dank für das Interesse an dieser fakultätsübergreifenden Thematik und für die intensive Begutachtung der Dissertation.

Zusätzlich gilt mein Dank Frau Prof. Dr. Gabriele M. König und Frau Prof. Dr. Ute Nöthlings für die Bereitschaft zur Teilnahme an der begutachtenden Promotionskommission.

Ein ganz herzlicher Dank geht an Herrn Prof. Dr. Michael Ludwig für seine großartige und unermüdliche Unterstützung. Er hatte immer ein offenes Ohr für meine Fragen und Probleme und brachte mich durch zahlreiche fachliche Diskussionen Schritt für Schritt weiter.

Ich möchte den im Projekt direkt involvierten CURE-Net-Teammitgliedern herzlich danken: Pia, Uerdingen, Alina C. Hilger, Franziska Kause, Florian Marsch, Greta Gross und Magdalena Schmidt.

Für die Durchführung der statistischen Auswertungen bedanke ich mich bei Herrn PD. Dr. Michael Knapp.

Bei allen Kollegen der Humangenetik im Life&Brain und BMZ bedanke ich mich herzlich für die gute Einarbeitung in diverse Labortechniken, die uneingeschränkte Unterstützung sowie hilfreiche wissenschaftliche Diskussionen und Ratschläge.

Mein Dank gilt den Projektpartnern für die gute Zusammenarbeit am 1. Anatomischen Institut, Universität Bonn (Herr Prof. Dr. Benjamin Odermatt und Frau Öznur Yilmaz); 2. Cologne Center for Genomics (CCG) an der Universität zu Köln (Herr Prof. Dr. Peter Nürnberg, Herr Dr. Holger Thiele und Frau Dr. Janine Altmüller); 3. Institut für Kardiovaskuläre Regeneration, Universität Frankfurt am Main (Herr PD Dr. Phillip Grote); 4. Abteilung für Entwicklungsgenetik, Medizinische Universität Wakayama, Japan (Herr Prof. Dr. Gen Yamada). Ich bedanke mich

ebenfalls für die gute Kooperation mit Herrn Prof. Dr. William Newman und Frau Dr. Glenda Beaman (Manchester Zentrum für Genomische Medizin, UK), und Frau Prof. Dr. Agneta Nordenskjöld (Karolinska Institut und Astrid Lindgren Kinderklinik, Stockholm, Schweden).

Nicht zuletzt möchte ich meiner Familie und meinen Freunden danken. Danke, dass ihr mich auf meinem Weg begleitet habt und mich stets unterstützt habt. Mein ganz besonderer und herzlicher Dank geht an meinen Eltern. Ich danke euch für eure uneingeschränkte und liebevolle Unterstützung.