

**Genomweite Identifizierung von kausalen
Mikrodeletionen oder -duplikationen
(Kopienzahlveränderungen) bei Patienten mit
angeborenen anorektalen Malformationen**

Inaugural-Dissertation

zur Erlangung des Doktorgrades

der Hohen Medizinischen Fakultät

der Rheinischen Friedrich-Wilhelms-Universität

Bonn

Julia Katharina Fabian

aus Freiburg im Breisgau

2024

Angefertigt mit der Genehmigung
der Medizinischen Fakultät der Universität Bonn

1. Gutachter: Prof. Dr. med. Heiko Martin Reutter
2. Gutachter: Prof. Dr. Dipl. Phys. Peter Krawitz

Tag der Mündlichen Prüfung: 15.10.2024

Aus dem Institut für Humangenetik
Direktor: Prof. Dr. med. Markus M. Nöthen

Inhaltsverzeichnis

	Abkürzungsverzeichnis	4
1.	Deutsche Zusammenfassung	6
1.1	Einleitung	6
1.2	Material und Methoden	7
1.3	Ergebnisse	9
1.4	Diskussion	14
1.5	Zusammenfassung	17
1.6	Literaturverzeichnis der deutschen Zusammenfassung	18
2.	Veröffentlichung	23
	Abstract	23
	Introduction	23
	Materials (Subjects) and Methods	24
	Results	24
	Discussion	27
	Conclusions	28
	References	28
3.	Danksagung	30

Abkürzungsverzeichnis

ARM	anorektale Malformation(en)
BAF	B-allele frequency
CNV	copy number variation (Kopienzahlveränderung)
DECIPHER	database of chromosomal imbalance and phenotype in humans using ensembl resources (Datenbank über chromosomales Ungleichgewicht und Phänotypen beim Menschen unter Verwendung von Ensemble-Ressourcen)
DGV	database of genomic variants (Datenbank genomischer Varianten)
DNA	desoxyribonucleic acid (Desoxyribonukleinsäure)
LRR	log R ratio
MAMEP	Molecular Anatomy of the Mouse Embryo Project (Molekulare Anatomie des Mausembryo-Projekts)
MGI	mouse genome informatics (Maus Genom Datenbank)
qPCR	quantitative polymerase chain reaction (quantitative Polymerase-Kettenreaktion)
SNP	single-nucleotide polymorphism (Einzelnukleotid-Polymorphismus)
Tab.	Tabelle
VATER/VACTERL	Vertebral defects (Wirbelkörperfehlbildungen) Anorectal malformations (anorektale Fehlbildungen) Cardiac defects (kardiale Fehlbildungen) Tracheoesophageal fistula (tracheo-ösophageale Fistel) Esophageal atresia (Ösophagusatresie) Renal malformations (renale Fehlbildungen)

Limb defects (Extremitätenfehlbildungen)

1. Deutsche Zusammenfassung

1.1 Einleitung

Das Krankheitsbild der Anorektalen Malformationen (ARM) beschreibt seltene Fehlbildungen des embryonalen Enddarms, bei denen es zu einem fehlgebildeten Analkanal und Rektum kommt. Das phänotypische Spektrum beinhaltet angeborene Analfehlbildungen mit und ohne Fisteln. Fisteln treten zwischen dem Rektum und verschiedenen Endpunkten, wie der Urethra, der Blase, dem Vestibulum vaginae, der Vagina und dem Perineum auf. Ohne Fistel auftretende ARM sind die Analatresie, die Rektumatresie, die Rektumstenose und die Kloakenpersistenz (Levitt *et al.*, 2007). In männlichen Neugeborenen findet sich die rektourethrale Fistel am häufigsten, bei weiblichen Neugeborenen überwiegt die rektovestibuläre Fistel (Wood *et al.*, 2018). ARM treten mit einer Prävalenz von 1 in 2.500 - 3.000 auf, wobei Jungen mit einem Faktor von 1,7 häufiger betroffen sind als Mädchen (Jenetzky, 2007). Es werden isolierte ARM von nicht-isolierten ARM unterschieden. Nicht-isolierte ARM sind definiert als das gemeinsame Auftreten einer ARM und weiteren angeborenen Fehlbildungen. Hierunter finden sich am häufigsten (zu ca. 80 %) urogenitale Begleitfehlbildungen (Stoll *et al.*, 2007). Nicht-isolierte ARM sind meist Teil definierter genetischer Syndrome (syndromale ARM) und machen etwa 60 % des gesamten Patientenspektrums der ARM aus. Die isolierte ARM tritt nicht-syndromal auf und steht nicht in Zusammenhang mit weiteren Fehlbildungen. Ca. 40 % aller ARM treten isoliert auf (Marcelis *et al.*, 2021). Je nach klinischer Ausprägung des Krankheitsbildes bedarf es bereits innerhalb der ersten Lebenstage des Neugeborenen einer chirurgischen Korrektur, die mit lebenslangen Einschränkungen, unter anderem mit Obstipationen, Stuhl- und Harn-Inkontinenz und dadurch psychischen Belastungen einhergehen kann (Levitt *et al.*, 2007). Die Ätiologie der ARM scheint multifaktoriell zu sein. Durch vorherige Studien konnte bereits nachgewiesen werden, dass einerseits verschiedenste Umweltfaktoren, andererseits genetische Faktoren eine wichtige Rolle spielen (Zwink *et al.*, 2011; van Rooij *et al.*, 2010; Schramm *et al.*, 2011; Wong *et al.*, 2013). Vor allem für syndromale Formen der ARM sind bereits diverse genetische Faktoren wie numerische oder strukturelle Chromosomenanomalien in der Literatur vorbeschrieben (Dworschak *et al.*, 2013;

Dworschak *et al.*, 2015), wohingegen für die isolierte ARM die Beteiligung von genetischen Risikofaktoren noch deutlich weniger erforscht ist. Unter anderem leisten Kopienzahlveränderungen (sog. copy number variations, CNVs) einen Beitrag zu der Entstehung einer ARM (Hilger *et al.*, 2013). Kopienzahlveränderungen sind Abweichungen der Anzahl an Kopien einzelner Abschnitte des Genoms. Die Größe der CNVs reicht von einzelnen bis zu mehreren Millionen Basenpaaren und umfasst sowohl Deletionen als auch Duplikationen. CNVs können vererbt werden (> 99 % der CNVs) oder *de novo* entstehen (< 1 % der CNVs) (McCarroll *et al.*, 2008; McCarroll, 2008). In der Normalbevölkerung selten vorkommende CNVs und *de novo* Varianten können der Grund für verschiedenste Erkrankungen sein. Ursächlich hierfür sind unterschiedliche Mechanismen. So kann ein CNV etwa innerhalb eines Gens liegen, dieses somit unterbrechen und dadurch zu einem Funktionsverlust führen. Des Weiteren können CNVs Gene, die eine Dosierungs-Sensitivität besitzen in ihrer Expressionsrate verändern und so zu einem klinischen Phänotyp führen. Ebenso können CNVs in intronischen regulatorischen Regionen liegen, wie beispielsweise Promotorregionen. Bei Unterbrechung dieser Regionen können CNVs somit Gene, die durch die entsprechenden Regionen beeinflusst werden, ebenfalls in ihrer Expressionsrate verändern (Zhang *et al.*, 2009).

Bereits seit einigen Jahren wird die Beteiligung von CNVs an der Entstehung der ARM erforscht. Die hier beschriebene Studie beschäftigte sich mit der Kausalität von CNVs an der Entwicklung einer ARM. Ziel der Studie war es, den Gesamtbeitrag von CNVs an der Expression von ARM weiter zu beleuchten und neue krankheitsverursachende CNVs und Gene, sogenannte Kandidatengene aufzufinden. Sie ist die bisher größte Studie ihrer Art.

1.2 Material und Methoden

An der Studie nahmen insgesamt 450 Patienten und ihre Eltern teil. Hiervon zeigten 212 Patienten eine syndromale und 238 Patienten eine isolierte ARM. Sie wurden über das bundesweite Netzwerk zu Congenitalen Uro-REktalen Fehlbildungen (CURE-Net, www.cure-net.de) rekrutiert. Die Patientenrekrutierung und Durchführung der genetischen

Studie genehmigte die Ethikkommission der Universität Bonn. Ferner wurden 4392 gesunde Kontrollpersonen in die Studie mit einbezogen, die ab dem Jahr 2000 an der Heinz Nixdorf RECALL (Risk Factors, Evaluation of Coronary Calcium and Lifestyle) Studie teilnahmen (Schmermund *et al.*, 2002).

Die DNA der Patienten und ihrer Eltern wurde mithilfe des Chemagic Magnetic Separation Modules I (Chemagen, Baesweiler, Deutschland) aus Blutproben oder unter Verwendung des Oragene DNA Kits (DNA Genotek Inc., ON, Kanada) aus Speichelproben isoliert. Aus der isolierten DNA erfolgte im Anschluss die Genotypisierung für die CNV-Untersuchung mittels single nucleotide polymorphism (SNP)-basiertem Array "Infinium Global Screening Array-24 v2.0" von Illumina (Illumina, San Diego, CA, USA), bei dem durch Hybridisierungen mit SNP-Fluoreszenzmarkersonden bekannte biallelische SNPs detektiert wurden. Aus den Daten des SNP-Arrays konnte anschließend mit Hilfe der Software PennCNV (Wang *et al.*, 2007) die Berechnung der CNVs durchgeführt werden. Die Software nutzt hierbei ein Hidden-Markov-Modell, das eine durch Duplikationen oder Deletionen erhöhte oder verminderte Intensität der SNP-Markersonden detektiert (sog. Log R Ratio, LRR) und das relative Verhältnis der Fluoreszenzsignale zwischen den jeweiligen Allelen (sog. B Allele Frequenz, BAF) berechnet. Somit kann PennCNV ermitteln, an welcher Position im Genom CNVs vorliegen. Hierauf folgte ein mehrschrittiges Filterverfahren, in dem zunächst die Qualität der Patienten- und Kontrollproben und im Anschluss die Qualität der einzelnen CNVs überprüft wurde.

Die Prüfung der Qualität der Patienten- und Kontrollproben erfolgte mittels der Faktoren Geschlechtsunstimmigkeiten, Call Rate und Standardabweichung. Dabei wurden Proben herausgefiltert, die eine Geschlechtsunstimmigkeiten zwischen den primär vorhandenen Angaben und den Informationen aus der Genotypisierung und eine reduzierte Call Rate von unter 98 % aufwiesen. Ebenso wurden Proben herausgefiltert, deren Anzahl an CNVs höher war, als das Doppelte der Standardabweichung der durchschnittlichen Anzahl an identifizierten CNVs pro Patient.

Die Qualität der einzelnen CNVs konnte mittels Log Bayes Faktor und der Anzahl an gefundenen SNPs pro CNV festgestellt werden. Die Qualitätskriterien wurden erreicht, wenn der Log Bayes Faktor eines CNVs oberhalb von 30 und die Anzahl an SNPs pro CNV oberhalb von drei lag. Durch einen nachfolgenden Vergleich aller CNVs aus den verbliebenen Patientenproben mit den CNVs der Kontrollproben konnten diejenigen

CNVs herausgefiltert werden, die in den gesunden Kontrollen vorhanden waren und somit als vermutlich nicht Phänotyp-assoziiert zu betrachten sind. Im nächsten Schritt wurden die CNVs in das Programm AnnotSV (Geoffroy *et al.*, 2021) eingespeist, das den CNVs bereits vorhandene Informationen aus verschiedenen Datenbanken zuordnet, wie beispielsweise Informationen aus der Datenbank genomischer Varianten (database of genomic variants, DGV). CNVs, die bereits in DGV beschrieben waren, wurden herausgefiltert, um nur CNVs zu identifizieren, die in der Bevölkerung noch nicht vorbeschrieben waren. Des Weiteren wurden CNVs herausgefiltert, die laut AnnotSV in nicht-codierenden Bereichen lagen und/oder keine Promotorregionen beinhalteten. Anschließend erfolgte eine visuelle Kontrolle der noch vorhandenen CNVs auf Marker und Signalqualität anhand der LRR und BAF mit dem Programm GenomeStudio Genotyping Module v2.0 und Illumina Genome Viewer version 2.0.4. Durch eine Literaturrecherche konnten CNVs priorisiert und anschließend mittels SYBR® Green basierter quantitativer Polymerase-Kettenreaktion (qPCR) validiert werden. Für die Literaturrecherche wurden die Datenbanken DECIPHER (<https://www.deciphergenomics.org>), gnomAD (<https://gnomad.broadinstitute.org>), MGI (<http://www.informatics.jax.org>), MAMEP (<http://mamep.molgen.mpg.de>), OMIM (<https://www.omim.org>), PubMed (<https://pubmed.ncbi.nlm.nih.gov>) und UCSC (<https://genome.ucsc.edu>) verwendet. Schlussendlich erfolgte die Identifizierung von Kandidatengenen unter Hinzunahme von drei Kriterien. Kandidatengene mussten in *de novo* aufgetretenen CNVs liegen, sollten also nicht vererbt worden sein. Dieses Kriterium konnte bei vorhandener elterlicher DNA durch die qPCR festgestellt werden. Des Weiteren mussten in der Literatur für die in Frage kommenden Gene Fehlbildungen vorbeschrieben sein, die mit anorektalen Fehlbildungen assoziiert waren. Als drittes Kriterium mussten die Gene im Mausembryo während der Embryonalstage 10,5 – 15,5 exprimiert sein, da dies die kritische Phase für die Bildung einer anorektalen Malformation in Mäusen darstellt (Kluth *et al.*, 2011).

1.3 Ergebnisse

Insgesamt wurden 450 Patienten genotypisiert. Durch die oben genannten Qualitätskriterien für die jeweiligen Proben wurden 54 Patienten aus der Studie ausgeschlossen. Zwei Patienten zeigten eine Geschlechtsunstimmigkeit, 40 Patienten

wiesen eine Call Rate unter 98 % auf und 12 Patienten besaßen mehr CNVs als das Doppelte der Standardabweichung. Insgesamt blieben somit 396 Patienten mit insgesamt 6,316 CNVs und 4,066 Kontrollen mit insgesamt 43,323 CNVs übrig. Die anschließenden Qualitätsfilter, die auf die einzelnen CNVs angewandt wurden (Log Bayes Faktor > 30, Anzahl an SNPs pro CNV > 3), schlossen 5,026 CNVs aus ARM Patienten und 3,2075 CNVs aus Kontrollproben aus. Die Suche nach CNVs, die sowohl in Patienten, als auch in Kontrollen vorhanden waren, ergab 887 CNVs, die herausgefiltert wurden. 117 Deletionen und 35 Duplikationen schieden aus, da sie laut AnnotSV in DGV vorbeschrieben waren, nicht in codierenden Bereichen lagen und/oder keine Promotorregionen beinhalten. Nach dem letzten Filterschritt, der visuellen Kontrolle, verblieben 82 Deletionen und 49 Duplikationen in der Studie. Aufgrund der Größe einiger CNVs war bereits vor den o.g. Filterschritten eindeutig, dass es sich bei diesen um echte CNVs handelte, sodass diejenigen CNVs aus den Filterschritten ausgeschlossen und nicht zusätzlich mittels qPCR validiert wurden. Schlussendlich fanden sich 14 Patienten mit CNVs, die priorisiert wurden. Da sich unter den Patienten ein Geschwisterpaar mit dem gleichen CNV befand, handelte es sich somit insgesamt um 13 CNVs (Tab. 1). Unter den 13 CNVs befanden sich zwei höchstwahrscheinlich unbalancierte Translokationen [46,XX,der(11)t(11;16)(q24.2;q22.2) und 46,XX,der(22)t(22;9)(q11.21;pter)], zwei terminale Duplikationen [46,XXdup(3)(q26.31q29) und 46,XXdup(17)(q25.3qter)], sieben weitere Deletionen (del2p13.2, del4p16.2, del7q31.33, del9p24.1, del16q12.1, del18q32, del22q11.21) und zwei weitere Duplikationen (dup2p13.2, dup17q12). Beide terminalen Duplikationen waren *de novo* aufgetreten, unter den weiteren Deletionen und Duplikationen befanden sich fünf vererbte CNVs, ein CNV mit unklarem Vererbungsstatus aufgrund fehlender elterlicher DNA und drei weitere *de novo* CNVs. In drei der insgesamt 13 CNVs konnten schließlich Kandidatengene (*LPP*, *FOXK2* und *SALL3*) identifiziert werden. Im Folgenden werden alle 14 Patienten beschrieben, in denen die 13 CNVs gefunden wurden.

Unbalancierte Translokationen

Beide Patienten mit unbalancierten Translokationen zeigten eine nicht-isolierte Form der ARM. Für alle Regionen der unbalancierten chromosomalen Translokationen waren in der

Literatur bereits Syndrome vorbeschrieben, die neben anderen Fehlbildungen einen ARM-Phänotypen aufführten.

Terminale *de novo* Duplikationen

Beide terminalen *de novo* Duplikationen waren in der Literatur noch nicht mit ARM assoziiert. Die *de novo* duplizierte Region 3q26.31-q29 war 25,1 Mb groß und beinhaltete neben anderen Genen das Entwicklungsgen *LPP*. Der Patient zeigte klinisch eine isolierte ARM, genauere Angaben über den Phänotyp des Patienten konnten jedoch nicht gemacht werden, da hierzu keine Daten vorlagen. Die zweite mikroskopische chromosomale *de novo* Duplikation wurde in der 0,48 Mb großen Region 17q25.3-qter gefunden. Innerhalb der Region fand sich unter anderem das Gen *FO XK2*. Der Patient wies ebenfalls eine isolierte ARM in Form einer perinealen Fistel auf.

Vererbte CNVs

Unter den vererbten CNVs fand sich zweimal die Region 2p13.2. Einerseits konnte in einer Patientin mit vestibulärer Fistel eine 0,31 Mb lange Deletion nachgewiesen werden, andererseits in einem Geschwisterpaar mit perinealen Fisteln eine 0,13 Mb lange Duplikation der gleichen Region. Einer der beiden Brüder wies zusätzlich eine Epispadie auf, der andere ein Hodenanhängsel und Taubheit im rechten Ohr. In beiden Familien war die Variante mittels qPCR nachweislich mütterlicherseits vererbt worden. Die Region 2p13.2 umfasst das Entwicklungsgen *EXOC6B*, das bisher mit dem Krankheitsbild der Osteochondrodysplasia spondyloepiphysaria (Campos-Xavier *et al.*, 2018) assoziiert war und in der Literatur noch keine vorbeschriebene Assoziation mit ARM hatte.

Ein weiterer vererbter CNV befand sich in Region 4p16.2. Bei dem CNV handelte es sich um eine Deletion, die paternal vererbt worden war. Der Patient wies klinisch eine rektoprostatistische Fistel, einen Morbus Hirschsprung und das Goldenhar-Syndrom auf. Die Deletion beinhaltete das Gen *STK23B*, das mit dem Ellis-van Creveld Syndrom assoziiert ist (Temtamy *et al.*, 2008). Das Ellis-van Creveld Syndrom ist eine chondrale und ektodermale Dysplasie, bei der es unter anderem zu genitalen Fehlbildungen wie der Epi- oder Hypospadie kommen kann (Kamal *et al.*, 2013).

Zwei weitere Regionen mit vererbten CNVs waren die paternal vererbte Deletion 7q31.33 bei einem Patienten mit syndromaler ARM mit VACTERL-Assoziation, die aufgrund

fehlender Daten jedoch nicht genauer definiert war und die maternal vererbte Duplikation 17q1 bei einem Patienten mit isolierter ARM mit perinealer Fistel. Die Deletion mit unklarem Vererbungsstatus lag in der Region 9q24.1, der Patient hatte eine syndromale ARM mit angeborenem Herzfehler, rechtsseitiger Nierenagenesie und Kryptorchismus.

Weitere *de novo* CNVs

Drei CNVs besaßen einen *de novo* Status. Zum einen lag bei einem Patienten mit isolierter ARM eine *de novo* Deletion in der Region 16q12.1 vor. Zum anderen besaß ein weiterer Patient eine 0,81 Mb große *de novo* Deletion von 18q32, die das Gen *SALL3* umfasste. Dieser Patient zeigte eine syndromale ARM mit rektoprostatistischer Fistel, Doppelnieren mit Megaureter und einem Maldescensus testis rechts.

Die *de novo* Deletion der Region 22q11.21 zeigte sich bei einem einzelnen Patienten und war ebenfalls Teil einer der beiden unbalancierten Translokationen (s.o.). Das sogenannte 22q11.2 Deletionssyndrom oder DiGeorge Syndrom ist bereits in der Literatur mit ARM assoziiert (Marcelis *et al.*, 2011) Der Patient präsentierte sich mit einer perinealen Fistel, einem offenen Ductus arteriosus, einem Ventrikelseptumdefekt, einer Pulmonalklappenstenose, einer Aryknorpel-Luxation und einem Bauchwandhernie.

Letztlich ließen sich aufgrund der weiter oben genannten Kriterien *de novo* Status, vorbeschriebene mit ARM assoziierte Fehlbildungen und Expression der Gene während der Embryonalstage 10,5 – 15,5 in Mäusen insgesamt drei Kandidatengene priorisieren. Hierbei handelte es sich um die Gene *LPP*, *FOXK2* und *SALL3*.

Tab. 1: Priorisierte CNVs

Die Kandidatengene sind hervorgehoben.

Der = Derivat, Dup = Duplikation, Del = Deletion, Mb = Megabasen, CNV = copy number variation

CNVs	Länge	Struktur-variante	Gen(e) innerhalb des CNVs	Vererbungsmuster
46,XX,der(11)t(11;16)(q24.2;q22.2)	7,8 Mb	Deletion	> 20 Gene	Unbekannt
	18 Mb	Duplikation	> 20 Gene	
46,XX,der(22)t(22;9)(q11.21;p)	2,5 Mb	Deletion	> 20 Gene	Unbekannt
	38.8 Mb	Duplikation	> 20 Gene	
46,XY,dup(3)(q26.31q29)	25,1 Mb	Duplikation	LPP und > 20 weitere	De novo
46,XY,dup(17)(q25.3qter)	0,48 Mb	Duplikation	FN3KRP, FOXK2, LOC101929552, MIR4525, RAB40B, WDR45B, FN3K, TBCD, ZNF750, B3GNTL1, TBCD	De novo
del2p13.2	0,31 Mb	Deletion	EXOC6B	Vererbt
dup2p13.2	0,13 Mb	Duplikation	EXOC6B	Vererbt
del4p16.2	0,26 Mb	Deletion	STK32B	Vererbt
del7q31.33	0,04 Mb	Deletion	GRM8	Vererbt
dup17q12	0,18 Mb	Duplikation	AP2B1, LINC02001, LOC107985033, PEX12, SLFN12L, SLFN14, SNORD7	Vererbt
del9p24.1	0,03 Mb	Deletion	JAK2	Unbekannt
del16q12.1	1,38 Mb	Deletion	ADCY7, HEATR3, MIR6771, PAPD5, BRD7, CYLD, LINC02127, LINC02168, LINC02178, LOC101927272, MIR3181, NKD1, NOD2, SNX20, SALL1	De novo

del18q32	0,81 Mb	Deletion	SALL3, ATP9B, NFATC1	<i>De novo</i>
del22q11.21	2,58 Mb	Deletion	> 20 Gene	<i>De novo</i>

1.4 Diskussion

Für die Regionen, in denen die zwei unbalancierten Translokationen [46,XX,der(11)t(11;16)(q24.2;q22.2) und 46,XX,der(22)t(22;9)(q11.21;pter)] lagen, sowie für die *de novo* Deletion (del22q11.21), die gleichermaßen in einer der beiden unbalancierten Translokationen vorkam, existierten bereits in der Literatur vorbeschriebene Fälle von Patienten mit ARM. Der Zusammenhang zwischen anorektalen Malformationen und diesen Regionen konnten hiermit weiter bestätigt werden. Beide terminalen *de novo* Duplikationen [46,XXdup(3)(q26.31q29) und 46,XXdup(17)(q25.3qter)] und alle weiteren CNVs (del2p13.2, del4p16.2, del7q31.33, del9p24.1, del16q12.1, del18q32, dup2p13.2 und dup17q12) waren dagegen noch nicht in der Literatur vorbeschrieben.

Die Gene in den fünf vererbten CNVs und in dem CNV mit unbekanntem Vererbungsstatus (del2p13.2, dup2p13.2, del4p16.2, del7q31.33, dup17q12, del9p24.1) wurden aufgrund ihres Vererbungsstatus nicht als potenzielle neue Kandidatengene priorisiert. Dennoch gab es Hinweise darauf, dass in zwei CNVs jeweils ein Gen möglicherweise krankheitsverursachende Effekte auf die Entstehung einer ARM haben könnte. Hierbei handelte es sich einerseits um die Region 2p13.2, die in einer Patientin deletiert und in einem Geschwisterpaar dupliziert war. Der CNV umfasste nur das Entwicklungsgen *EXOC6B*. Mittels der online verfügbaren Maus Genom Datenbank (mouse genome informatics, MGI) konnte eine starke Expression von *exoc6b* im viszeralen Organsystem der Maus während der Embryonalstage 10,5 – 18 nachgewiesen werden. Da die kritische Phase für die Entwicklung einer ARM in Mäusen in den Embryonaltagen 10,5 – 15,5 liegt, stimmt die Expressionsrate von *EXOC6B* somit zeitlich mit dieser überein. Weil sich in den 450 Patienten gleich drei Patienten mit Anomalien in dieser Region fanden, legt dies nahe, dass für *EXOC6B* eine gewisse Dosierungs-Sensitivität bestehen könnte.

Der zweite CNV, der Hinweise auf krankheitsverursachende Effekte im Rahmen einer ARM lieferte, war die Deletion der Region 4p16.2, die das Gen *STK32B* umfasste.

Strukturvariationen in 4p16.2 sind mit dem autosomal-rezessiven Ellis-van Creveld Syndrom assoziiert, bei dem es unter anderem zu urogenitalen Fehlbildungen wie der Epi- oder Hypospadie kommen kann. Urogenitale Fehlbildungen sind eng mit nicht-isolierten ARM assoziiert (Stoll *et al.*, 2007). Da neben dem Vererbungsstatus das Ellis-van Creveld Syndrom ein autosomal-rezessives und kein autosomal-dominantes Muster besitzt, kann hier lediglich ein gewisser Dosierungseffekt vermutet werden.

Neben *EXOC6B* und *STK32B*, die einen Hinweis auf eine Assoziation mit der Genese einer ARM lieferten, konnten drei weitere Kandidatengene definiert werden, die höchstwahrscheinlich ursächlich für die Entwicklung einer ARM sind. Bei diesen Genen handelte es sich um *LPP*, *FO XK2* und *SALL3*.

Das Kandidatengen *LPP* konnte in der Region 3q28 lokalisiert werden, die bei einem der 14 Patienten dupliziert war. Der gesamte CNV umspannte die Region 3q26.31 bis 3q29 und beinhaltete somit das gesamte Gen. *LPP* war bereits in vorigen Studien mit dem VATER/VACTERL-Spektrum assoziiert (Hernández-García *et al.*, 2012). Das VATER/VACTERL-Spektrum ist durch das gemeinsame Auftreten von mindestens drei der folgenden Fehlbildungen charakterisiert: Wirbelsäulendefekte (V), ARM (A), Herzfehler (C), tracheoösophageale Fisteln mit Ösophagusatresie (TE), Nierenfehlbildungen (R) und Fehlbildungen der Extremitäten (L) (Solomon, 2011). Ferner konnte mittels MAMEP (Molecular Anatomy of the Mouse Embryo Project, Molekulare Anatomie des Mausembryo-Projekts) in Mäusen eine Expression von *lpp* im Enddarm während der Embryonalstage 13,5 – 15 nachgewiesen werden, sodass in Zusammenschau ein Zusammenhang zwischen dem dieses Gen betreffenden CNV und der ARM des Patienten möglich erscheint.

Als weiteres Kandidatengen ließ sich das Entwicklungsgen *FO XK2* in der Duplikation der Region 17q25.3-qter identifizieren. Die Region 17q25.3-qter beinhaltet Teile von Exon 8 und das gesamte Exon 9 des Gens *FO XK2*. *FO XK2* wies im Mäuseembryo am Embryonaltag 15,5 eine starke Expressionsrate in den Harnstrukturen auf. Zudem existierte in der Literatur bereits für das FOX-Gencluster der Region 16q24 ein Zusammenhang mit Ösophagusatresien, die eine der assoziierten Fehlbildungen mit nicht-isolierten ARM darstellen (Gehlen *et al.*, 2022). Einige in der Literatur beschriebene Patienten mit Aberrationen des FOX-Genclusters zeigten zusätzlich anorektale

Malformationen wie Analstenosen (Stankiewicz *et al.*, 2009), sodass insgesamt von einem Dosierungseffekt auf die Bildung anorektaler Malformationen ausgegangen werden kann.

Das dritte Kandidatengen (*SALL3*) war durch die Deletion der Region 18q32 im Ganzen betroffen. Auch mit Genen aus der Familie von *SALL3* konnten bereits in vorigen Studien Assoziationen mit syndromalen ARM hergestellt werden. Hierbei handelte es sich um *SALL1* und *SALL4*. Genetische Veränderungen in *SALL1* verursachen das Townes-Brocks Syndrom, dessen Phänotyp unter anderem anorektale Malformationen beinhaltet (Borozdin *et al.*, 2006), Veränderungen in *SALL4* führen zum Duane-Radial Ray Syndrom, bei dem es ebenfalls zu anorektalen Malformationen kommen kann (van de Putte *et al.*, 2020). Neben den beiden durch *SALL1* und *SALL4* verursachten Syndromen konnten mittels DECIPHER (Database of Chromosomal Imbalance and Phenotype in Humans Using Ensembl Resources, Firth *et al.*, 2009) noch weitere mit ARM assoziierte Fehlbildungen in der Literatur aufgefunden werden, die direkt *SALL3* betrafen. So zeigten Patienten mit *de novo* Deletionen in *SALL3* unter anderem Phänotypen, die häufig mit ARM assoziiert sind, wie Vaginalatresien, Hypospadien, Nierenagenesien und vesikoureteraler Reflux (Stoll *et al.*, 2007). *Sall3* wird darüber hinaus in Mäusen während der Embryonalstage 7,5 – 8,75 schwach im Endoderm exprimiert, aus dem sich in späteren Embryonalphasen der Enddarm entwickelt. Insgesamt ist anzunehmen, dass *SALL3* an der Entstehung einer ARM beteiligt ist.

Die hier verwendete Methode der CNV-Analyse mittels SNP-Arrays ist einerseits eine standardisierte und häufig verwendete Methode in der humangenetischen Diagnostik, andererseits birgt sie jedoch gewisse Einschränkungen in der Erforschung kausaler CNVs. So nutzt die Methode bereits bekannte SNPs und berechnet anhand derer die Position und Länge der CNVs. Dadurch bleibt jedoch die genaue Bruchstelle der CNVs unbekannt. Somit ist nicht klar, ob beispielsweise die drei Patienten mit CNVs in der Region 2p13.2 (ein einzelner Patient und ein Geschwisterpaar) genau die gleichen Bruchstellen besitzen oder unterschiedliche. Zudem wurden in der Studie nur codierende Regionen betrachtet. Da in nicht-codierenden Bereichen jedoch regulatorische Regionen liegen, könnten Auswirkungen von CNVs auf diese Bereiche und somit weitere potenziell pathogene Strukturvarianten übersehen worden sein. Die Aussagekraft der hier definierten CNVs ist außerdem dadurch beschränkt, dass keine Informationen über die

Phänotypen der Eltern vorlagen, insbesondere derjenigen Elternteile, die die CNVs vererbt hatten. CNVs, die durch gesunde Elternteile vererbt wurden, haben höchstwahrscheinlich nicht zu der Entwicklung einer anorektalen Malformation beigetragen. Bei Erfüllen der anderen beiden Kriterien für Kandidatengene (vorbeschriebene mit ARM assoziierte Fehlbildungen und Expression der Gene während der Embryonalstage 10,5 – 15,5 in Mäusen) kann hierbei jedoch trotzdem ein gewisser Dosierungseffekt vermutet werden. Bei Eltern, die ebenfalls einen ARM-Phänotyp aufweisen, liegt jedoch die Vermutung nahe, dass die entsprechend vererbten aber ansonsten selten vorkommenden CNVs krankheitsverursachend wären.

Diese Studie identifizierte drei wahrscheinlich kausale CNVs (dup3q26.31-q29, dup17q25.3-qter und del18q32), die nur in jeweils einem Patienten vorkamen, vermutlich da es sich um äußerst seltene CNVs handelte. Obwohl diese Studie aktuell die größte ihrer Art ist, die die Kausalität von CNVs im Zusammenhang mit der Entstehung von anorektalen Malformationen untersucht, müssten noch weitaus größere Kohorten verwendet werden, um sehr seltene CNVs in mehr als einem Patienten zu bestätigen.

1.5 Zusammenfassung

Die hier beschriebene CNV-Analyse konnte verdeutlichen, wie wichtig die Rolle von CNVs in der Ätiologie der ARM ist. Darüber hinaus wurden neue krankheitsverursachende CNVs und Kandidatengene gefunden und somit ein weiterer Beitrag zur Erforschung der genetischen Ursachen der ARM geleistet. Insgesamt zeigten sich in den 450 Patienten mit anorektalen Malformationen 14 Patienten mit mutmaßlich pathogenen CNVs. Dies entspricht einem Anteil von ungefähr 3 %. In vorherigen Studien entsprach der Anteil an Patienten mit krankheitsverursachenden CNVs ungefähr 5 %, sodass sich hier eine annäherungsweise Übereinstimmung ergab (Marcelis *et al.*, 2011; Hilger *et al.*, 2013). Die vorliegende Studie ist nach derzeitigem Wissensstand die aktuell größte Studie, die die Kausalität von CNVs im Zusammenhang mit der Entwicklung einer ARM untersucht. ARM stellen ein schwerwiegendes Krankheitsbild dar, das mit teilweise lebenslangen Einschränkungen einhergeht. Die aktuelle Studie bestätigte die Annahme, dass CNVs einen Beitrag zur Expression von ARM leisten und konnte neue krankheitsverursachende CNVs und Kandidatengene aufzeigen. Perspektivisch wird die weitere systematische

Identifizierung von potenziell pathogenen CNVs und Genen dazu beitragen, die zugrundeliegenden molekularen Prozesse noch umfangreicher zu verstehen. So kann eine verbesserte genetische Beratung bezüglich erblicher Hintergründe betroffener Familien gewährleistet und die Wahrscheinlichkeit des Wiederholungsrisikos bei Kindern besser eingeschätzt werden.

1.6 Literaturverzeichnis der deutschen Zusammenfassung

Borozdin W, Steinmann K, Albrecht B, Bottani A, Devriendt K, Leipoldt M, Kohlhasse J. Detection of heterozygous SALL1 deletions by quantitative real time PCR proves the contribution of a SALL1 dosage effect in the pathogenesis of Townes-Brocks syndrome. *Human mutation*. 2006. 27: 211-212

Campos-Xavier B, Rogers RC, Niel-Bütschi F, Ferreira C, Unger S, Spranger J, Superti-Furga A. Confirmation of spondylo-epi-metaphyseal dysplasia with joint laxity, EXOC6B type. *American journal of medical genetics*. 2018. 176: 2934-2935

Dworschak GC, Draaken M, Marcelis C, de Blaauw I, Pfundt R, van Rooij IA, Bartels E, Hilger A, Jenetzky E, Schmiedeke E, Grasshoff-Derr S, Schmidt D, Märzheuser S, Hosie S, Weih S, Holland-Cunz S, Palta M, Leonhardt J, Schäfer M, Kujath C, Rissmann A, Nöthen MM, Zwink N, Ludwig M, Reutter H. 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. *American journal of medical genetics*. 2013. 161A: 3035-3041

Dworschak GC, Draaken M, Hilger AC, Schramm C, Bartels E, Schmiedeke E, Grasshoff-Derr S, Märzheuser S, Holland-Cunz S, Lacher M, Jenetzky E, Zwink N, Schmidt D, Nöthen MM, Ludwig M, Reutter H. Genome-wide mapping of copy number variations in patients with both anorectal malformations and central nervous system abnormalities. *Birth defects research. Part A, Clinical and molecular teratology*. 2015. 103: 235-242

Firth HV, Richards SM, Bevan AP, Clayton S, Corpas M, Rajan D, Van Vooren S, Moreau Y, Pettett RM, Carter NP. DECIPHER: Database of Chromosomal Imbalance and Phenotype in Humans Using Ensembl Resources. *American journal of human genetics*. 2009. 84: 524-533

Gehlen J, Giel AS, Köllges R, Haas SL, Zhang R, Trcka J, Sungur AÖ, Renziehausen F, Bornholdt D, Jung D, Hoyer PD, Nordenskjöld A, Tibboel D, Vlot J, Spaander MCW, Smigiel R, Patkowski D, Roeleveld N, van Rooij IA, de Blaauw I, Hölscher A, Pauly M, Leutner A, Fuchs J, Niethammer J, Melissari MT, Jenetzky E, Zwink N, Thiele H, Hilger AC, Hess T, Trautmann J, Marks M, Baumgarten M, Bläss G, Landén M, Fundin B, Bulik CM, Pennimpede T, Ludwig M, Ludwig KU, Mangold E, Heilmann-Heimbach S, Moebus S, Herrmann BG, Alsabeah K, Burgos CM, Lilja HE, Azodi S, Stenström P, Arnbjörnsson E, Frybova B, Lebensztejn DM, Debek W, Kolodziejczyk E, Kozera K, Kierkus J, Kaliciński P, Stefanowicz M, Socha-Banasiak A, Kolejwa M, Piaseczna-Piotrowska A, Czkwianianc E, Nöthen MM, Grote P, Rygl M, Reinshagen K, Spychalski N, Ludwikowski B, Hubertus J, Heydweiller A, Ure B, Muensterer OJ, Aubert O, Gosemann JH, Lacher M, Degenhardt P, Boemers TM, Mokrowiecka A, Małecka-Panas E, Wöhr M, Knapp M, Seitz G, de Klein A, Oracz G, Brosens E, Reutter H, Schumacher J. First genome-wide association study of esophageal atresia identifies three genetic risk loci at CTNNA3, FOXF1/FOXC2/FOXL1, and HNF1B. *HGG advances*. 2022. 3: 100093

Geoffroy V, Guignard T, Kress A, Gaillard JB, Solli-Nowlan T, Schalk A, Gatinois V, Dollfus H, Scheidecker S, Muller J. AnnotSV and knotAnnotSV: a web server for human structural variations annotations, ranking and analysis. *Nucleic acids research*. 2021. 49: W21-W28

Hernández-García A, Brosens E, Zaveri HP, de Jong EM, Yu Z, Namwanje M, Mayle A, Fernandes CJ, Lee B, Blazo M, Lalani SR, Tibboel D, de Klein A, Scott DA. Contribution of LPP copy number and sequence changes to esophageal atresia, tracheoesophageal fistula, and VACTERL association. *American journal of medical genetics*. 2012. 158A: 1785-1787

Hilger A, Schramm C, Pennimpede T, Wittler L, Dworschak GC, Bartels E, Engels H, Zink AM, Degenhardt F, Müller AM, Schmiedeke E, Grasshoff-Derr S, Märzheuser S, Hosie S, Holland-Cunz S, Wijers CH, Marcelis CL, van Rooij IA, Hildebrandt F, Herrmann BG, Nöthen MM, Ludwig M, Reutter H, Draaken M. De novo microduplications at 1q41, 2q37.3, and 8q24.3 in patients with VATER/VACTERL association. *European journal of human genetics: EJHG*. 2013. 21: 1377-1382

Jenetzky E. Prevalence estimation of anorectal malformations using German diagnosis related groups system. *Pediatric surgery international*. 2007. 23: 1161-1165

Kamal R, Dahiya P, Kaur S, Bhardwaj R, Chaudhary K. Ellis-van Creveld syndrome: A rare clinical entity. *Journal of oral and maxillofacial pathology: JOMFP*. 2013. 17: 132-135

Kluth D, Fiegel HC, Metzger R. Embryology of the hindgut. *Seminars in pediatric surgery*. 2011. 20: 152-160

Levitt MA, Peña A. Anorectal malformations. *Orphanet journal of rare diseases*. 2007. 2: 33

Marcelis C, de Blaauw I, Brunner H. Chromosomal anomalies in the etiology of anorectal malformations: a review. *American journal of medical genetics*. 2011. 155A: 2692-2704

Marcelis C, Dworschak G, de Blaauw I, van Rooij I. Genetic Counseling and Diagnostics in Anorectal Malformation. *European journal of pediatric surgery : official journal of Austrian Association of Pediatric Surgery*. 2021. 31: 482-491

McCarroll SA. Extending genome-wide association studies to copy-number variation. *Human molecular genetics*. 2008. 17: R135-142

McCarroll SA, Kuruvilla FG, Korn JM, Cawley S, Nemesh J, Wysoker A, Shapero MH, de Bakker PI, Maller JB, Kirby A, Elliott AL, Parkin M, Hubbell E, Webster T, Mei R, Veitch J, Collins PJ, Handsaker R, Lincoln S, Nizzari M, Blume J, Jones KW, Rava R, Daly MJ, Gabriel SB, Altshuler D. Integrated detection and population-genetic analysis of SNPs and copy number variation. *Nature genetics*. 2008. 40: 1166-1174

Schramm C, Draaken M, Bartels E, Boemers TM, Schmiedeke E, Grasshoff-Derr S, Märzheuser S, Hosie S, Holland-Cunz S, Baudisch F, Priebe L, Hoffmann P, Zink AM, Engels H, Brockschmidt FF, Aretz S, Nöthen MM, Ludwig M, Reutter H. De novo duplication of 18p11.21-18q12.1 in a female with anorectal malformation. *American journal of medical genetics*. 2011. 155A: 445-449

Schmermund A, Möhlenkamp S, Stang A, Grönemeyer D, Seibel R, Hirche H, Mann K, Siffert W, Lauterbach K, Siegrist J, Jöckel K H, Erbel R. 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. *American heart journal*. 2002. 144: 212–218

Solomon BD. VACTERL/VATER Association. *Orphanet journal of rare diseases*. 2011. 6: 56

Stankiewicz P, Sen P, Bhatt SS, Storer M, Xia Z, Bejjani BA, Ou Z, Wiszniewska J, Driscoll DJ, Maisenbacher MK, Bolivar J, Bauer M, Zackai EH, McDonald-McGinn D, Nowaczyk MM, Murray M, Hustead V, Mascotti K, Schultz R, Hallam L, McRae D, Nicholson AG, Newbury R, Durham-O'Donnell J, Knight G, Kini U, Shaikh TH, Martin V, Tyreman M, Simonic I, Willatt L, Paterson J, Mehta S, Rajan D, Fitzgerald T, Gribble S, Prigmore E, Patel A, Shaffer LG, Carter NP, Cheung SW, Langston C, Shaw-Smith C. Genomic and genic deletions of the FOX gene cluster on 16q24.1 and inactivating mutations of FOXF1 cause alveolar capillary dysplasia and other malformations. *American journal of human genetics*, 2009. 84: 780–791

Stoll C, Alembik Y, Dott B, Roth MP. Associated malformations in patients with anorectal anomalies. *European journal of medical genetics*. 2007. 50: 281-290

Temtamy SA, Aglan MS, Valencia M, Cocchi G, Pacheco M, Ashour AM, Amr KS, Helmy SM, El-Gammal MA, Wright M, Lapunzina P, Goodship JA, Ruiz-Perez VL. Long interspersed nuclear element-1 (LINE1)-mediated deletion of EVC, EVC2, C4orf6, and STK32B in Ellis-van Creveld syndrome with borderline intelligence. *Human mutation*. 2008. 29: 931-938

van de Putte R, Dworschak GC, Brosens E, Reutter HM, Marcelis CLM, Acuna-Hidalgo R, Kurtas NE, Steehouwer M, Dunwoodie SL, Schmiedeke E, Märzheuser S, Schwarzer N, Brooks AS, de Klein A, Sloots CEJ, Tibboel D, Brisighelli G, Morandi A, Bedeschi MF, Bates MD, Levitt MA, Peña A, de Blaauw I, Roeleveld N, Brunner HG, van Rooij IALM, Hoischen A. A Genetics-First Approach Revealed Monogenic Disorders in Patients With ARM and VACTERL Anomalies. *Frontiers in pediatrics*. 2020. 8: 310

van Rooij IA, Wijers CH, Rieu PN, Hendriks HS, Brouwers MM, Knoers NV, de Blaauw I, Roeleveld N. Maternal and paternal risk factors for anorectal malformations: a Dutch case-control study. *Birth defects research. Part A, Clinical and molecular teratology*. 2010. 88: 152-158

Wang K, Li M, Hadley D, Liu R, Glessner J, Grant SF, Hakonarson H, Bucan M. PennCNV: an integrated hidden Markov model designed for high-resolution copy number variation detection in whole-genome SNP genotyping data. *Genome research*. 2007. 17: 1665-1674

Wong EH, Cui L, Ng CL, Tang CS, Liu XL, So MT, Yip BH, Cheng G, Zhang R, Tang WK, Yang W, Lau YL, Baum L, Kwan P, Sun LD, Zuo XB, Ren YQ, Yin XY, Miao XP, Liu J, Lui VC, Ngan ES, Yuan ZW, Zhang SW, Xia J, Wang H, Sun XB, Wang R, Chang T, Chan IH, Chung PH, Zhang XJ, Wong KK, Cherny SS, Sham PC, Tam PK, Garcia-Barcelo MM. Genome-wide copy number variation study in anorectal malformations. *Human molecular genetics*. 2013. 22: 621-631

Wood RJ, Levitt MA. Anorectal Malformations. *Clinics in colon and rectal surgery*. 2018. 31: vol. 31: 61-70

Zhang F, Gu W, Hurles ME, Lupski JR. Copy number variation in human health, disease, and evolution. *Annual review of genomics and human genetics*. 2009. 10: 451–481

Zwink N, Jenetzky E, Brenner H. Parental risk factors and anorectal malformations: systematic review and meta-analysis. *Orphanet journal of rare diseases*. 2011. 6: 25

2. Veröffentlichung


www.nature.com/ejhg
ARTICLE **OPEN**


Genome-wide identification of disease-causing copy number variations in 450 individuals with anorectal malformations

Julia Fabian¹, Gabriel C. Dworschak^{1,2,3}, Lea Waffenschmidt¹, Luca Schierbaum¹, Charlotte Bendixen^{1,4}, Stefanie Heilmann-Heimbach¹, Sugirthan Sivalingam^{5,6,7}, Andreas Buness^{5,6,7}, Nicole Schwarzer⁸, Thomas M. Boemers⁹, Eberhard Schmiedeke¹⁰, Jörg Naser¹¹, Johannes Leonhardt¹², Ferdinand Kosch¹³, Sandra Weih¹⁴, Helen Maya Gielen¹⁵, Stuart Hosie¹⁶, Carmen Kabs¹⁶, Markus Palta¹⁷, Stefanie Märzheuser¹⁸, Lena Marie Bode¹⁹, Martin Lacher¹⁹, Frank-Mattias Schäfer²⁰, Maximilian Stehr²⁰, Christian Knorr²¹, Benno Ure²², Katharina Kleine²³, Udo Rolle²⁴, Marcin Zaniew²⁵, Grote Phillip²⁶, Nadine Zwink²⁷, Ekkehart Jenetzky^{27,28}, Heiko Reutter^{1,29} and Alina C. Hilger^{1,30,31}✉

© The Author(s) 2022

Anorectal malformations (ARM) represent a spectrum of rare malformations originating from a perturbed development of the embryonic hindgut. Approximately 60% occur as a part of a defined genetic syndrome or within the spectrum of additional congenital anomalies. Rare copy number variations (CNVs) have been associated with both syndromic and non-syndromic forms. The present study represents the largest study to date to explore the contribution of CNVs to the expression of ARMs. SNP-array-based molecular karyotyping was applied in 450 individuals with ARM and 4392 healthy controls. CNVs were identified from raw intensity data using PennCNV. Overlapping CNVs between cases and controls were discarded. Remaining CNVs were filtered using a stringent filter algorithm of nine filter steps. Prioritized CNVs were confirmed using qPCR. Filtering prioritized and qPCR confirmed four microscopic chromosomal anomalies and nine submicroscopic CNVs comprising seven microdeletions (del2p13.2, del4p16.2, del7q31.33, del9p24.1, del16q12.1, del18q32, del22q11.21) and two microduplications (dup2p13.2, dup17q12) in 14 individuals (12 singletons and one affected sib-pair). Within these CNVs, based on their embryonic expression data and function, we suggest *FOXP2*, *LPP*, and *SALL3* as putative candidate genes. Overall, our CNV analysis identified putative microscopic and submicroscopic chromosomal rearrangements in 3% of cases. Functional characterization and re-sequencing of suggested candidate genes is warranted.

European Journal of Human Genetics; <https://doi.org/10.1038/s41431-022-01216-5>

INTRODUCTION

Anorectal malformations (ARM) represent a spectrum of rare malformations originating from a perturbed development of the embryonic hindgut. ARM present with a male to female ratio of 1.7, the overall birth prevalence has been estimated with 1 in 2500–3000 live births [1]. Approximately 60% occur as a part of a

defined genetic syndrome or within the spectrum of additional congenital anomalies, the remaining 40% present isolated non-syndromic [2]. The current knowledge of the underlying molecular mechanisms of isolated non-syndromic ARM is still limited. Based on observations such as reduced reproduction rate one can assume a significant fraction of underlying de novo events.

¹Institute of Human Genetics, Medical Faculty of the University Bonn & University Hospital Bonn, Bonn, Germany. ²Department of Neuropediatrics, University Hospital Bonn, Bonn, Germany. ³Institute of Anatomy, Medical Faculty, University of Bonn, Bonn, Germany. ⁴Unit of Pediatric Surgery, Department of General, Visceral, Vascular and Thoracic Surgery, University Hospital Bonn, Bonn, Germany. ⁵Institute for Medical Biometry, Informatics and Epidemiology, Medical Faculty, University of Bonn, Bonn, Germany. ⁶Institute for Genomic Statistics and Bioinformatics, Medical Faculty, University of Bonn, Bonn, Germany. ⁷Core Unit for Bioinformatics Data Analysis, Medical Faculty, University of Bonn, Bonn, Germany. ⁸SoMA, The German Patient Support Organization for Anorectal Malformations and Hirschsprung Disease, Munich, Germany. ⁹Department of Pediatric Surgery and Pediatric Urology, Children's Hospital of Cologne Amsterdamer Strasse, Cologne, Germany. ¹⁰Clinic for Pediatric Surgery and Pediatric Urology, Klinikum Bremen Mitte, Bremen, Germany. ¹¹Department of Pediatric Surgery, General Hospital, Chemnitz, Germany. ¹²Department of Pediatric Surgery, Children's Hospital Braunschweig, Braunschweig, Germany. ¹³Department of Pediatric Surgery, Städtisches Klinikum Karlsruhe, Karlsruhe, Germany. ¹⁴Department of Pediatric Surgery, Medical Center, Faculty of Medicine, University of Freiburg, Freiburg, Germany. ¹⁵Department of Pediatric Surgery, Asklepios Klinik Nord Heidberg, Hamburg, Deutschland. ¹⁶Muenchen Klinik gGmbH, Muenchen, Klinik Schwabing, Technische Universität München, Munich, Germany. ¹⁷Department of Pediatric Surgery, Evangelisches Krankenhaus Hamm, Hamm, Germany. ¹⁸Department of Pediatric Surgery, Rostock University Medical Center, Rostock, Germany. ¹⁹Department of Pediatric Surgery, University of Leipzig, Leipzig, Germany. ²⁰Department of Pediatric Surgery and Pediatric Urology, Knopfsche Kinderklinik-Klinik Hallerwiese, Nürnberg, Germany. ²¹Department of Pediatric Surgery and Orthopedics, University Children's Hospital Regensburg (KUNO) at the Hospital St. Hedwig of the Order of St. John, Regensburg, Germany. ²²Center of Pediatric Surgery Hannover, Hannover Medical School, Hannover, Germany. ²³Department of Pediatric Surgery, Evangelisches Krankenhaus Oberhausen, Oberhausen, Germany. ²⁴Department of Pediatric Surgery and Pediatric Urology, Goethe University Frankfurt, Frankfurt, Germany. ²⁵Department of Pediatrics, University of Zielona Góra, Zielona Góra, Poland. ²⁶Institute of Cardiovascular Regeneration, Center for Molecular Medicine, University of Frankfurt, Frankfurt am Main, Germany. ²⁷Department of Child and Adolescent Psychiatry, University Medical Center of the Johannes Gutenberg University Mainz, Mainz, Germany. ²⁸Faculty of Health, School of Medicine, University of Witten/Herdecke, Witten, Germany. ²⁹Division of Neonatology and Pediatric Intensive Care, Department of Pediatrics and Adolescent Medicine, Friedrich-Alexander University Nürnberg-Erlangen, Erlangen, Germany. ³⁰Department of Pediatrics and Adolescent Medicine, Friedrich-Alexander University Nürnberg-Erlangen, Erlangen, Germany. ³¹Research Center On Rare Kidney Diseases (RECORD), University Hospital Erlangen, 91054 Erlangen, Germany. ✉email: heiko.reutter@uk-erlangen.de; alina.hilger@uk-erlangen.de

Received: 18 March 2022 Revised: 3 July 2022 Accepted: 6 October 2022
Published online: 01 November 2022

SPRINGER NATURE

The experience in medical genetics suggests that such mutational events will comprise genomic alterations of different size ranging from small changes affecting single nucleotides to large alterations resulting in losses or gains of several thousand to millions of base-pairs. To test this hypothesis, we previously, between 2011 and 2017, systematically employed array-based molecular karyotyping in altogether 224 individuals with syndromic and non-syndromic ARM [3–10]. In total, we identified a pathogenic de novo CNV in 5 % of cases. 11 cases presented with syndromic ARM according to Rasmussen et al. [11]. Accordingly, Wong et al. using molecular karyotyping in 363 Han Chinese with sporadic ARM and 4006 controls found an enrichment of rare long duplications among individuals with syndromic ARM [12]. The aims of our present study were (i) to explore the overall contribution of CNVs to the expression of ARMs, (ii) to identify novel disease-causing CNVs, and (iii) to identify novel candidate genes.

MATERIALS (SUBJECTS) AND METHODS

Cases and controls

Affected individuals and their families were recruited within the framework of the German “Network for Congenital Uro-Rectal malformations” (www.cure-net.de). All experiments involving human materials adhered to the recommendations of the World Medical Association, as stated in the revised Declaration of Helsinki from Seoul 2008. The proposed project has been approved by the Ethics Commission of the University of Bonn. The study sample encompassed 450 individuals with ARM. A total of 212 individuals presented with a syndromic, 238 with a non-syndromic form. Furthermore, we included 4392 healthy controls [13]. For 351 affected individuals, the DNA of both parents was available, for 57 affected individuals, only maternal, and for 5 affected individuals, only paternal DNA was available. For 37 affected individuals no parental sample was available.

DNA isolation and array-based molecular karyotyping

Genomic DNA of patients and their parents was isolated from the blood samples by using the Chemagic Magnetic Separation Module I (Chemagen, Baesweiler, Germany), or from saliva samples using the Oragene DNA Kit (DNA Genotek Inc., ON, Canada). Cases and Controls were genotyped using the single nucleotide polymorphism (SNP)-based array “Infinium Global Screening Array-24 v2.0”. This array contains >600,000 SNPs that are evenly distributed across the genome. All chromosomal positions are given in human genome build hg19 (GRCh37 Genome Reference Consortium Human Build 37 (GRCh37)).

CNV filtering

CNV calling was accomplished using PennCNV. The algorithm of PennCNV uses a Hidden-Markov model to estimate putative CNVs from log R ratio (normalized intensity data) and B allele frequency (allele frequency data) of each SNP. Quality criteria used to exclude individuals due to failed genotyping were (I) a gender mismatch between assumed and called gender; (II) a call rate <98%; (III) a number of CNVs per affected individual that exceeded double of the standard deviation of the mean number of CNVs called per individual. Subsequent criteria to exclude single CNVs were (IV) maximum log Bayes factor below 30; and (V) if the CNV comprised less than 3 consecutive called SNPs. These steps were followed by filtering of the remaining CNVs for prioritization of potential disease-causing CNVs. Filter steps included (VI) comparison of CNVs in cases and inhouse controls. All CNVs identified in cases that were overlapped by one or more CNV form controls were not further considered potential disease causing. Beforehand, all 4394 inhouse controls had been quality checked accordingly using criteria (I)–(V). For the remaining CNVs the information for each CNV from PennCNV was combined with the information from the annotation and ranking tool AnnotSV, and filtered out in a manual filter step (VII) [14]. Here, CNVs comprising

deletions were discarded when overlapping with deletion CNVs annotated in the Database of Genomic Variants (DGV) and CNVs comprising duplications were discarded when overlapping with duplication CNVs annotated in the Database of Genomic Variants (DGV). Furthermore, (VIII) CNVs that did not reside within coding or promotor regions were filtered out. (IX) Using GenomeStudio Genotyping Module v2.0 and the Illumina Genome Viewer version 2.0.4 the remaining CNVs were visually inspected. (X) Final prioritization of the remaining CNV regions was performed by gathered information from publicly available databases: Gnomad (<https://gnomad.broadinstitute.org>), OMIM (<https://www.omim.org>), UCSC (<https://genome.ucsc.edu>), MGI (<http://www.informatics.jax.org>), and PubMed (<https://pubmed.ncbi.nlm.nih.gov>). Genome build hg19/GRCh37 was used for all data.

CNV confirmation

To confirm all prioritized CNVs and to identify de novo events when parental DNA was available, quantitative polymerase chain reaction (qPCR) was carried out. SYBR Green (Applied Biosystems, Foster City, CA) was used for qPCR reactions. The evaluation of the copy numbers was performed using the comparative CT method ($\Delta\Delta C_t$ method). Three housekeeping genes were used to compare relative copy counts (*CFTR*, *BNC1*, *RNA*). The primer sequences are shown in Supplementary Table 1.

RESULTS

CNV analysis

After the application of quality filter step I (gender mismatch), two patients were excluded. Through step II (call rate <98%), 40 patients were discarded and after step III (exceeded double of standard deviation), 12 more patients were excluded. In the remaining 396 individuals with ARM, a total of 6316 CNVs, were called. Within the remaining 4066 controls, a total of 43,323 CNVs were called. Regarding the quality filter for each CNV, filter steps (IV) and (V) led to the exclusion of 5026 CNVs among ARM individuals and the exclusion of 32,075 CNVs among healthy control individuals. Comparison of frequencies in cases and controls (VI) excluded additional 887 CNVs. In this filter step, we identified three individuals with trisomy 21 (not further mentioned). Application of filter steps VII and VIII filtered out 117 microdeletions and 35 microduplications. Filter step IX yielded a final count of 82 microdeletions and 49 microduplications. According to the annotation of the respective CNVs in gnomAD (<https://gnomad.broadinstitute.org>), OMIM (<https://www.omim.org>), UCSC (<https://genome.ucsc.edu>), MGI (<http://www.informatics.jax.org>), and PubMed (<https://pubmed.ncbi.nlm.nih.gov>), we ultimately prioritized and qPCR confirmed four microscopic chromosomal anomalies and nine submicroscopic CNVs comprising seven microdeletions (del2p13.2, del4p16.2, del7q31.33, del9p24.1, del16q12.1, del18q32, del22q11.21) and two microduplications (dup2p13.2, dup17q12).

Unbalanced chromosomal translocations

Two affected individuals carried unbalanced terminal translocations including 46,XX,der(11)t(11;16)(q24.2;q22.2) and 46,XX,der(22)t(22;9)(q11.21;pter) (Table 1).

Individual 1 presented with an ARM in form of a perineal fistula, congenital intraventricular hemorrhage with hydrocephalus, bilateral clubfoot, dysplastic pelvic kidney (left), bilateral atresia of submandibular gland, left sided aplasia of the thyroid gland, choanal atresia, hearing disability, and macular delocalization. Her chromosomal anomaly comprised an unbalanced translocation 46,XX,der(11)t(11;16)(q24.2;q22.2) with terminal deletion 11q24.2-qter. The first deleted SNP is located at chromosomal position chr11:127,078,525 and the last deleted SNP at chr11:134,923,456. Deletion of 11q23.3-qter has been associated with Jacobsen syndrome (JBS, MIM #147791). JBS is a contiguous gene deletion syndrome involving terminal chromosome 11q. Key phenotypic features are pre- and postnatal growth retardation, intellectual

Table 1. Unbalanced chromosomal translocations and microscopic chromosomal de novo duplications.

	Individual 1	Individual 2	Individual 3	Individual 4
CNV	46,XX,der(11)t(11;16) (q24.2;q22.2)	46,XX,der(22)t(22;9) (q11.21;p)	46,XY,dup(3) (q26.31-q29)	46,XY,dup(17) (q25.3-qter)
Type of ARM	Perineal fistula	Perineal fistula	–	Perineal fistula
Phenotypes				
Heart	–	–	–	–
CNS	Intraventricular hemorrhage of the newborn	–	–	–
Esophagus	–	–	–	–
Extremities	Bilateral clubfoot	–	–	–
Vertebral bodies	–	–	–	–
Hips	–	–	–	–
Kidney	Dysplastic pelvic kidney (left)	Kidney hypoplasia (left). Bilateral hydronephrosis	–	–
Urinary tract	–	Persistent urogenital sinus	–	–
Genital organs	–	–	–	–
Other	Bilateral atresia of submandibular gland. Left sided aplasia of thyroid. Choanal atresia. Hearing disability. Macular delocalization	Premature infant (36 + 0), small for gestation age (2090 g)	–	–

ARM anorectal malformation, CNS central nervous system.

disability (ID), characteristic facial dysmorphism, thrombo- or pancytopenia. Further clinical features are congenital heart defects (CHD), congenital anomalies of the kidneys and urinary tract (CAKUT), gastrointestinal tract, genitalia, central nervous system (CNS) anomalies and/or skeleton. Previously, Mattina et al. reported gastrointestinal tract malformations in 18–25% of individuals with JBS including pyloric stenosis, ARM, and less frequently, duodenal atresia, annular pancreas, or gut malrotation [15]. Besides terminal deletion 11q24.2-qter Individual 1 carried also an 18 Mb duplication of chromosomal region 16q22.2-qter. The first duplicated SNP is located at chromosomal position chr16:71.956.505, the last duplicated SNP at chromosomal position chr16:90.161.959. Key features described for partial trisomy 16qter (ORPHA:96106) include low birth weight, failure to thrive, hypotonia, ID, CHD, limb anomalies and joint contractures, facial dysmorphism with a high prominent forehead, down slanting and small palpebral fissures, hypertelorism, periorbital edema, low-set and abnormal ears, prominent nose, micrognathia, CAKUT, genital anomalies, and ARM. Interestingly, several reports in the literature associated dup16pterqter with ARM [16].

Individual 2 presented with ARM in form of a perineal fistula, kidney hypoplasia (left), bilateral hydronephrosis, persistent urogenital sinus, and small for gestational age. Her chromosomal anomaly comprised an unbalanced translocation with a 2.5 Mb deletion of chromosomal region 22q11.21. The first deleted SNP was located at chromosomal position chr22:18.875.445 and the last deleted SNP was located at chromosomal position chr22:21.461.607. The 22q11.2 deletion syndrome is the most common microdeletion syndrome. It is not surprising that several cases of ARM have been described among patients with this deletion [16] (see below Individual 14). Individual 2 also carried a 38.8 Mb duplication of 9p. The first duplicated SNP was located at chromosomal position chr9:209.325, the last duplicated SNP was located at chromosomal position chr9:39.021.035. Common features of a 9p trisomy include ID, craniofacial dysmorphisms, skeletal alterations, CNS anomalies, CHD, and less common CAKUT. While ARM has not been associated with trisomy 9p, several reports in the literature mention ARM in association with mosaic or non-mosaic complete trisomy 9 [16].

Microscopic chromosomal de novo duplications

Two affected individuals carried each a terminal duplication including 46,XXdup(3)(q26.31-q29) and 46,XXdup(17)(q25.3-qter) (Table 1).

Individual 3 presented with an isolated ARM. It is uncertain if he presented with additional minor stigmata. His chromosomal anomaly comprised a 25.1 Mb duplication of chromosome 3q26.31-q29 (Fig. 1a). The first duplicated SNP was located at chromosomal position chr3:171.615.261, the last at chromosomal position chr3:196.805.528. Common features of 3q duplication syndrome are a hirsutism, synophrys, broad nasal root, anteverted nares, downturned corners of the mouth, malformed ears, CHD, CAKUT and genital anomalies, ID, and growth retardation. To the best of our knowledge, this is the first report of duplication 3q26.31-q29 in an individual with ARM.

Individual 4 presented with an isolated ARM in form of a perineal fistula. He did not present with any other congenital anomaly or stigmata. His chromosomal anomaly comprised an 0.48 Mb duplication of 17q25.3-qter (Fig. 1b). The first duplicated SNP was located at chromosomal position chr17:80.545.076, the last at chromosomal position chr17:81.033.874. To the best of our knowledge, this is the first report of duplication of 17q25.3-qter in an individual with ARM.

Rare inherited CNVs and CNVs of unknown inheritance

Six affected individuals carried very rare inherited submicroscopic CNVs (Table 2). All of these CNVs were absent in 4066 healthy inhouse controls (CNV frequency <0.0003) and DGV.

Individual 5 presented with an isolated ARM in form of a recto-vestibular fistula. Besides her ARM, she had no co-occurring anomalies. She carried a maternally inherited 316.074 bp deletion of chromosomal region 2p13.2. The first deleted SNP was located on chromosomal position Chr2:72.623.204, the last at chromosomal position Chr2:72.939.279. Within this deletion resides the developmental gene *EXOC6B*. *EXOC6B* has been associated with autosomal-recessive inherited spondyloepimetaphyseal dysplasia with joint laxity, type 3 (MIM #618395).

Individuals 6 and 7 are brothers who both presented with ARM in form of a perineal fistula. In addition, Individual 6 presented with a testicular appendage and right-sided deafness. His brother presented

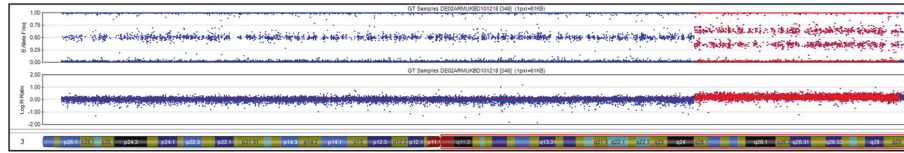
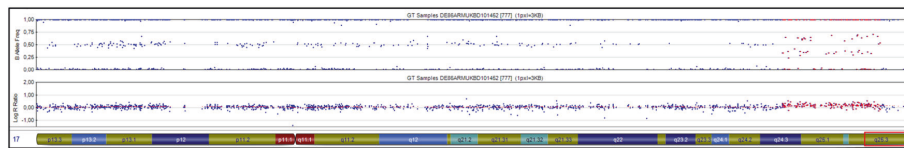
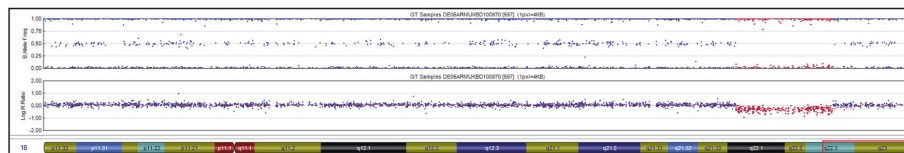
a) Individual 3: 25,1 Mb duplication of chromosome 3q26.31-q29 including gene *LPP*.b) Individual 4: 0,48 Mb duplication of 17q25.3qter including gene *FOXP2*.c) Individual 13: 0,81 Mb de novo deletion of 18q32 including gene *SALL3*.

Fig. 1 CNVs harboring putative ARM candidate genes.

with additional epispadias. Both carried a maternally inherited 132,652 bp duplication of chromosomal region 2p13.2. The first duplicated SNP was located at chromosomal position Chr2:72.882.934, the last at chromosomal position Chr2:73.015.587. Within this duplication resides the above-mentioned developmental gene *EXOC6B*. While the above *EXOC6B*-associated spondyloepimetaphyseal dysplasia with joint laxity, type 3 does not present with any visceral organ affection, the findings in individual 5 and in the two brothers 6 and 7 suggests a possible involvement of *EXOC6B* or the respective genomic region in ARM formation.

Individual 8 presented with ARM in form of a rectoprostatic fistula, aganglionosis of the Meissner's plexus (submucosa, M. Hirschsprung) and craniofacial Goldenhar syndrome. He carried a paternally inherited 265,097 bp deletion of chromosomal region 4p16.2. The first deleted SNP was located at chromosomal position Chr4:5.162.593, the last SNP at chromosomal position Chr4:5.427.691. Within this deletion resides the developmental gene *STK32B*, which has been associated with Ellis-van Creveld syndrome, an autosomal-recessive skeletal dysplasia with co-occurring genital anomalies, e.g., hypospadias.

Individual 9 presented with an unclassified ARM as part of her VATER/VACTERL association, with further phenotypic data missing. She carried a paternally inherited 47,083 bp deletion of chromosomal region 7q31.33. The first deleted SNP was located at chromosomal position Chr7:126.392.369, the last at chromosomal position Chr7:126.439.453. Within this deletion resides the developmental gene *GRM8*, so far associated with any human disease.

Individual 10 presented with ARM in form of a perineal fistula. He carried a 181,925 bp maternally inherited duplication of chromosomal region 17q1. The first duplicated SNP was located at chromosomal position Chr17:33.786.176, the last duplication at chromosomal position Chr17:33.968.102. Within this duplication reside five coding genes comprising *AP2B*, *PEX12*, *SLFN12L*, *SLFN14*, and *SNORD7* and two long non-coding RNAs LINC02001 and LOC107985033. Of the coding genes, *PEX12* has been associated

with the autosomal-recessive Peroxisome biogenesis disorder 3A (Zellweger, MIM #614859). *SLFN14* has been associated with the autosomal-dominant "Bleeding disorder, platelet-type, 20 (MIM #616913)".

Individual 11 presented with ARM in form of a recto-vesical fistula as part of his VATER/VACTERL association. He presented with additional CHD, right-sided kidney agenesis, and cryptorchidism. His chromosomal anomaly comprised a 27,823 bp deletion of chromosomal region 9q24.1 (Table 2). The first deleted SNP was located on chromosomal position Chr9:49.73.080, the last at chromosomal position Chr9:5.000.904. Here, we had no parental samples available for testing and can therefore not conclude, if the CNV has occurred de novo or has been transmitted from a healthy parent. Within the deletion resides the developmental gene *JAK2*, a gene that has been associated with broad spectrum of human hematologic malignancies (MIM *147796) but not with congenital visceral anomalies.

Rare submicroscopic de novo CNVs

In total we detected and confirmed three submicroscopic de novo CNVs (Table 3).

Individual 12 presented with an ARM in form of a perineal fistula and right-sided auricular hypoplasia. She carried a 1.38 Mb de novo deletion of chromosomal region 16q12.1. The first deleted SNP was located at chromosomal position chr16:50.135.837, the last at chromosomal position chr16:51.522.044. Within the deleted region reside nine coding genes including *ADCY7*, *HEATR3*, *PAPDS5*, *BRD7*, *CYLD*, *NKD1*, *NOD2*, *SNX20*, and *SALL1*. Furthermore, the deletion comprised two microRNAs (MIR6771 and MIR3181) and four long non-coding RNAs (LINC02127, LINC02168, LINC02178, LOC101927272).

Individual 13 presented with ARM in form of a rectoprostatic fistula, duplicated kidney with megaureter, and right-sided maldescensus testis. He carried a 0.81 Mb de novo deletion of chromosomal region 18q32 (Fig. 1c). The first deleted SNP was located at chromosomal position chr18:76.501.085, the last at

Table 2. Rare inherited CNVs and CNVs of unknown inheritance.

Individuals	Inherited CNVs	First and last deleted SNP	Size	Type	Genes within CNV
Individual 5	del2p13.2	Chr2:72.623.204-72.939.279	316,076 bp	Deletion	<i>EXOC6B</i>
Individuals 6 and 7	dup2p13.2	Chr2:72.882.934-73.015.587	132,654 bp	Duplication	<i>EXOC6B</i>
Individual 8	del4p16.2	Chr4:5.162.593-5.427.691	265,099 bp	Deletion	<i>STK32B</i>
Individual 9	del7q31.33	Chr7:126.392.369-126.439.453	47,085 bp	Deletion	<i>GRM8</i>
Individual 10	dup17q12	Chr17:33.786.176-33.968.102	181,927 bp	Duplication	<i>AP2B1, LINC02001, LOC107985033, PEX12, SLFN12L, SLFN14, SNORD7</i>
<i>CNV of unknown inheritance</i>					
Individual 11	del9p24.1	Chr9:4.973.080-5.000.904	27,825 bp	Deletion	<i>JAK2</i>

Individual 10 presented with ARM in form of a perineal fistula. He carried a 181.925 bp maternally inherited duplication of chromosomal region dup17q1.

chromosomal position 18:77.373.296. Within the deletion reside three coding genes comprising *SALL3*, *ATP9B*, and *NFATC1*.

Individual 14 presented with an ARM in form of a perineal fistula. In addition, she presented with patent ductus arteriosus, ventricular septal defect and pulmonary valve stenosis. Furthermore, she had luxation of the arytenoid cartilages and an abdominal wall hernia. Her chromosomal anomaly comprised an isolated 2.58 Mb deletion of 22q11.21 (see above Individual 2). The first deleted SNP was located on chromosomal position Chr22:18.875.445, the last at chromosomal position Chr22:21.461.607. DECIPHER database (Database of genomic variation and Phenotype in Humans using Ensembl Resources) [17] lists 1267 CNVs of chromosomal region 22q11.2. These comprise 667 heterozygous deletions. Within these 667 deletions are 14 deletions that have been described in association with ARM. The smallest region of overlap of all these deletions comprises 1.39 Mb. Here, further genetic studies are warranted to prioritize a possible ARM associated human disease-gene.

DISCUSSION

Congenital ARM represent a rare but severe spectrum of birth defects of the hindgut, with life-long impairments. The current knowledge of the underlying molecular mechanisms is still limited.

The primary aim of our present study was to identify novel highly penetrant structural genomic variants at the size of CNVs, and to describe in-depth the overall contribution of highly penetrant CNVs to the expression of the ARM spectrum. Among 450 ARM individuals, we identified three individuals with trisomy 21, two unbalanced chromosomal translocations, and two microscopic large de novo duplications (Table 1). Furthermore, we detected and confirmed five very rare, putatively novel submicroscopic CNVs in six affected individuals (four independent individuals and one sib-pair) (Table 2). None of these CNVs were found in any of our 4,066 healthy inhouse controls (CNV frequency <0.0003) nor in DGV. To the best of our knowledge, none of these inherited CNVs have been previously associated with human ARM phenotypes or other human diseases (PubMed search February 2022). Furthermore, we detected and confirmed three submicroscopic de novo CNVs in three independent ARM individuals. Our results confirm our own previous observations in which about 5% of affected individuals were found to carry putative disease-causing CNVs.

All large microscopic chromosomal rearrangements detected here comprise previously syndromic ARM associated regions. Within the rare inherited CNVs, we were unable to identify any candidate gene that would have stood out due to its function or association with known human ARM related disease. However, our findings in Individual 5 and the brothers Individual 6 and 7 suggest a possible dosage effect of *EXOC6B* to be involved in human ARM formation. Biallelic recessive variants in *EXOC6B* have been associated with spondyloepimetaphyseal dysplasia with joint

laxity, type 3, which does not present with visceral organ affection. However, according to Mouse Genome Database (MGD) *Exoc6b* shows strong expression in the mouse visceral organ system during embryonic development from embryonic day (ED) 10.5–18, including the critical embryonic timeframe for ARM organ formation [18]. Furthermore, we detected a heterozygous deletion in Individual 8 comprising *STK32B*, which has been associated with Ellis-van Creveld syndrome, an autosomal-recessive skeletal dysplasia with co-occurring genital anomalies, e.g., hypospadias. Hence, *STK32B* dosage effects might as well be implicated in ARM formation as assumed for *EXOC6B*, especially since we did not explore the respective trans alleles for disease-causing variants on the level of single base-pairs.

In five individuals we detected microscopic and submicroscopic chromosomal de novo rearrangements in which we prioritized three putative candidate genes. In Individual 12 we detected a 1.38 Mb de novo deletion of chromosomal region 16q12.1 (Table 3). Within the deleted region reside nine coding genes including *ADCY7*, *HEATR3*, *PAPDS*, *BRD7*, *CYLD*, *NKD1*, *NOD2*, *SNX20*, and *SALL1*. Previously, heterozygous deletions spanning *SALL1* have been associated with Townes-Brocks syndrome (MIM # 107480) [19]. Key phenotypic features of Townes-Brocks syndrome include microcephaly, auricular anomalies such as microtia, CHD, ARM, hypospadias, CAKUT, skeletal anomalies of the hands and feet. Hence, the present genetic finding and the individual's phenotype suggest the diagnosis of Townes-Brocks syndrome in Individual 12.

In Individual 13 we detected and confirmed a 0.81 Mb de novo deletion of chromosomal region 18q32 (Table 3 and Fig. 1c). Within the deletion reside three coding genes comprising *SALL3*, *ATP9B*, and *NFATC1*. Previously, the two developmental genes *SALL1* and *SALL4* have been associated with syndromic ARM [19, 20]. Furthermore, DECIPHER database [17] lists 47 individuals with heterozygous de novo losses of the *SALL3* comprising genomic region. Two individuals presented with vaginal atresia, five with hypospadias, three with renal agenesis, and two with vesicoureteral reflux. Moreover, according to MAMEP database (<http://mamep.molgen.mpg.de/index.php>), *Sall3* is expressed in the early mouse embryo (Theiler stage, TS15) in distinct structures. The highest expression is detected in the neural tube along the complete body axis and in the caudal end mesenchyme. Additional expression domains are detected in the developing limb buds and the brain area. In younger TS12 embryos, weaker staining is also detected in the endoderm areas, from which the hindgut will develop [21]. Together with these previous findings and the here observed findings in Individual 13, we suggest a dosage effect of *SALL3* to be involved in the formation of human ARM or uro-rectal malformations in a broader spectrum. In Individual 3 we detected a 25.1 Mb duplication of chromosome 3q26.31-q29 (Fig. 1a). Within this large duplication resides among other genes the developmental gene *LPP*. Genetic variants of different size in *LPP* have been previously associated with the VATER/VACTERL association [22–25]. The acronym VATER/VACTERL association (OMIM #192350)

Table 3. Patients with de novo CNVs.

	Individual 12	Individual 13	Individual 14
CNV	del16q12.1	del18q32	del22q11.21
Type of ARM	Perineal fistula	Rectoprostic fistula	Perineal fistula
Phenotypes			
Heart			PDA, VSD, pulmonary valve stenosis
CNS	–	–	–
Esophagus	–	–	–
Extremities	–	–	–
Vertebral bodies	–	–	–
Hips	–	–	–
Kidney	–	Duplex kidney with megaureter	–
Urinary tract	–	–	–
Genital organs	–	Maldescensus testis (right)	–
Other	Auricular hypoplasia (right)	–	Luxation of the arytenoid cartilages, Abdominal wall hernia

ARM anorectal malformation, CNS central nervous system, PDA patent ductus arteriosus, VSD ventricular septal defect, NA not available.

refers to the non-random co-occurrence of the following component features (CFs): vertebral defects (V), anorectal malformations (ARM) (A), cardiac defects (C), tracheoesophageal fistula with or without esophageal atresia (TE), renal malformations (R), and limb defects (L) [24]. Individuals may present with additional anomalies; however, the clinical diagnosis requires the presence of at least three CFs [22]. These previously associated dosage effects of *LPP* include an approximate 451,000 bp heterozygous loss spanning nucleotides 189,395,885–189,951,376 on chromosome 3q (nucleotide position based on hg18) comprising only *LPP* gene [22]. According to MAMEP database *Lpp* is expressed widespread in TS15 embryos with distinct expression domains in the developing somites and the mesenchyme [21]. Additional expression domains are detected in the gut and in older TS17 embryos expression can be also detected in the hindgut area [21]. Together, we suggest *LPP* as a putative ARM candidate gene. In Individual 4 we detected a 0.48 Mb duplication of 17q25.3-qter (Fig. 1b). Within this duplicated region resides the developmental gene *FO XK2*. Very recently, the *FOX*-gene-cluster on human chromosome 16q24 has been associated with multifactorial esophageal atresia, the most common congenital intestinal atresia of the upper intestinal tract [26]. Furthermore, *Foxk2* has a widespread expression pattern in the TS15 stage mouse embryo with distinct expression domains in the branchial arches, developing limbs and the epithelial structure of the forebrain and midbrain [27]. Moreover, *Foxk2* is strongly expressed in almost all developing urinary tract structures at mouse embryonic day 15.5 with distinct expression patterns in mouse embryonic kidney [28]. Taken together, our findings and previous observations suggest dosage effects of *FO XK2* to be implicated in human ARM formation. Our focus on CNVs comprising coding regions limited our current and previous analysis. Thereby we might have missed deep intronic small CNVs affecting only regulatory regions such as TATA boxes or TAD boundaries, possibly implicated in the formation of human ARM [29]. An additional limitation of our study was the method we carried out for CNV calling. For molecular karyotyping we employed an SNP-based array containing >600,000 SNPs distributed across the genome, followed by the estimation of the CNV-positions through PennCNV based on gains or losses of the SNPs. As a result, the precise CNV breakpoints remain unknown. We therefore cannot determine exactly, if patients with the apparent same CNV such as patients 5, 6 and 7 share the exact same breakpoints.

Further studies on our generated SNP-based array datasets, which were clearly beyond the scope of our present study might elucidate the complete contribution of CNVs to the formation of human ARM and explain larger proportions of its unknown genetic background.

CONCLUSIONS

This large scale CNV analysis with a focus on high penetrant chromosomal rearrangements in 450 individuals with ARM identified putative disease-causing microscopic and submicroscopic chromosomal rearrangements in 3% of cases. Within these CNVs reside putative candidate genes *FO XK2*, *LPP*, and *SALL3*. The confirmation of the putative candidate genes as ARM disease-genes will lead to new diagnostic possibilities, provide families and clinicians with greater knowledge of the causes of the disorder, and allow a precise estimation of a recurrence risk.

DATA AVAILABILITY

CNV data that support the findings of this study have been deposited in "ClinVar-Database" [<https://www.ncbi.nlm.nih.gov/clinvar/>] with the accession codes SCV002576529–SCV002576543.

REFERENCES

- Jenetzky E. Prevalence estimation of anorectal malformations using German diagnosis related groups system. *Pediatr Surg Int*. 2007;23:1161–5.
- Marcelis C, Dworschak G, de Blaauw I, van Rooij IALM. Genetic counseling and diagnostics in anorectal malformation. *Eur J Pediatr Surg*. 2021;31:482–91.
- Schramm C, Draaken M, Bartels E, Boemers TM, Aretz S, Brockschmidt FF, et al. De novo microduplication at 22q11.21 in a patient with VACTERL association. *Eur J Med Genet*. 2011;54:9–13.
- Schramm C, Draaken M, Bartels E, Boemers TM, Schmiedeke E, Grasshoff-Derr S, et al. De novo duplication of 18p11.21-18q12.1 in a female with anorectal malformation. *Am J Med Genet A*. 2011;155A:445–9.
- Baudisch F, Draaken M, Bartels E, Schmiedeke E, Bagci S, Bartmann P, et al. CNV analysis in monozygotic twin pairs discordant for urorectal malformations. *Twin Res Hum Genet*. 2013;16:802–7.
- Dworschak GC, Draaken M, Marcelis C, de Blaauw I, Pfundt R, van Rooij IALM, et al. 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*. 2013;161A:3035–41.
- Hilger A, Schramm C, Pennimpede T, Wittler L, Dworschak GC, Bartels E, et al. De novo microduplications at 1q41, 2q37.3, and 8q24.3 in patients with VATER/VACTERL association. *Eur J Hum Genet*. 2013;21:1377–82.

8. Dworschak GC, Draaken M, Hilger AC, Schramm C, Bartels E, Schmiedeke E, et al. Genome-wide mapping of copy number variations in patients with both anorectal malformations and central nervous system abnormalities. *Birth Defects Res A Clin Mol Teratol.* 2015;103:235–42.
9. Dworschak GC, Crétolle C, Hilger A, Engels H, Korsch E, Reutter H, et al. Comprehensive review of the duplication 3q syndrome and report of a patient with Currarino syndrome and de novo duplication 3q26.32-q27.2. *Clin Genet.* 2017;91:661–71.
10. Zhang R, Marsch F, Kause F, Degenhardt F, Schmiedeke E, Märzheuser S, et al. Array-based molecular karyotyping in 115 VATER/VACTERL and VATER/VACTERL-like patients identifies disease-causing copy number variations. *Birth Defects Res.* 2017;109:1063–9.
11. Rasmussen SA, Olney RS, Holmes LB, Lin AE, Keppler-Noreuil KM, Moore CA, et al. Guidelines for case classification for the National Birth Defects Prevention Study. *Birth Defects Res A Clin Mol Teratol.* 2003;67:193–201.
12. Wong EH, Cui L, Ng CL, Tang CS, Liu XL, So MT, et al. Genome-wide copy number variation study in anorectal malformations. *Hum Mol Genet.* 2013;22:621–31.
13. Schmermund A, Möhlenkamp S, Stang A, Grönemeyer D, Seibel R, Hirche H, et al. 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.* 2002;144:212–8.
14. Geoffroy V, Guignard T, Kress A, Gaillard JB, Solli-Nowlan T, Schalk A, et al. AnnotSV and knotAnnotSV: a web server for human structural variations annotations, ranking and analysis. *Nucleic Acids Res.* 2021;49:W21–W28. W1
15. Mattina T, Perrotta CS, Grossfeld P. Jacobsen syndrome. *Orphanet J Rare Dis.* 2009;7:9.
16. Marcellis C, de Blaauw I, Brunner H. Chromosomal anomalies in the etiology of anorectal malformations: a review. *Am J Med Genet A.* 2011;155A:2692–704.
17. Firth HV, Richards SM, Bevan AP, Clayton S, Corpas M, Rajan D, et al. DECIPHER: Database of Chromosomal Imbalance and Phenotype in Humans Using Ensembl Resources. *Am J Hum Genet.* 2009;84:524–33.
18. Bult CJ, Blake JA, Smith CL, Kadin JA, Richardson JE, et al. Mouse Genome Database (MGD) 2019. *Nucleic Acids Res.* 2019;47:D801–D806. the Mouse Genome Database Group. 2019D1
19. Borozdin W, Steinmann K, Albrecht B, Bottani A, Devriendt K, Leipoldt M, et al. Detection of heterozygous SALL1 deletions by quantitative real time PCR proves the contribution of a SALL1 dosage effect in the pathogenesis of Townes-Brooks syndrome. *Hum Mutat.* 2006;27:211–2.
20. Al-Baradie R, Yamada K, St Hilaire C, Chan WM, Andrews C, McIntosh N, et al. Duane radial ray syndrome (Okiihiro syndrome) maps to 20q13 and results from mutations in SALL4, a new member of the SAL family. *Am J Hum Genet.* 2002;71:1195–9.
21. Geffers L, Herrmann B, Eichele G. Web-based digital gene expression atlases for the mouse. *Mamm Genome.* 2012;23:525–38.
22. Arrington CB, Patel A, Bacino CA, Bowles NE. 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.* 2010;152A:2919–23.
23. Hernández-García A, Brosens E, Zaveri HP, de Jong EM, Yu Z, Namwanje M, et al. Contribution of LPP copy number and sequence changes to esophageal atresia, tracheoesophageal fistula, and VACTERL association. *Am J Med Genet A.* 2012;158A:1785–7.
24. Kolvenbach CM, van der Ven AT, Kause F, Shril S, Scala M, Connaughton DM, et al. Exome survey of individuals affected by VATER/VACTERL with renal phenotypes identifies phenocopies and novel candidate genes. *Am J Med Genet A.* 2021;185:3784–92.
25. Solomon BD. VACTERL/VATER association. *Orphanet J Rare Dis.* 2011;16:56.
26. Gehlen J, Giel AS, Köllges R, Haas SL, Zhang R, Trcka J, et al. First genome-wide association study of esophageal atresia with or without tracheoesophageal fistula (EA/TEF) identifies three genetic risk loci at CTNNA3, FOXF1/FOXC2/FOXL1 and HNF1B. *HHG Adv.* 2022;3:100093.
27. Harding SD, Armit C, Armstrong J, Brennan J, Cheng Y, Haggarty B, et al. The GUDMAP database – an online resource for genitourinary research. *Development.* 2011;138:2845–53.
28. Gray PA, Fu H, Luo P, Zhao Q, Yu J, Ferrari A, et al. Mouse brain organization revealed through direct genome-scale TF expression analysis. *Science.* 2004;306:2255–7.
29. Lifton RP, Goldberg ML, Karp RW, Hogness DS. The organization of the histone genes in *Drosophila melanogaster*: functional and evolutionary implications. *Cold Spring Harb Symposia Quant Biol.* 1978;42:1047–51.

ACKNOWLEDGEMENTS

We thank all patients and their families for their participation, as well as the German self-help organizations for individuals with anorectal malformations (SoMA e.V.) for their assistance with recruitment.

AUTHOR CONTRIBUTIONS

Conceptualization: HR and ACH. Methodology: JF, SHH, SS, AB, and GCD. Software: SS and AB. Validation: JF, CB, LS, and LW. Formal analysis: JF, SS, and ACH. Investigation: JF, GCD, HR, and ACH. Resources: HR, NS, TMB, ES, JN, JL, FK, SW, HMG, SH, CK, MP, SM, GG, ML, MS, MS, CK, BU, KK, UR, and MZ. Data curation: SS. Writing—original draft preparation: JF, ACH, PG, and HR. Writing—review and editing: all authors. Visualization: all authors. Supervision: HR and ACH. Funding acquisition: HR. All authors have read and agreed to the published version of the manuscript.

FUNDING

ACH was supported by BONFOR grant O-149.0123 and by the Else Kröner-Fresenius-Stiftung and the Eva Luise und Horst Köhler Stiftung – Project No: 2019_KollegSE04. HR was supported by the grants RE 1723/1-1 and RE 1723/1-3 from the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG). GCD has been supported by BONFOR grant O-120.0001 and by the Herbert-Reeck foundation (2019). We acknowledge financial support by Deutsche Forschungsgemeinschaft and Friedrich-Alexander-Universität Erlangen-Nürnberg within the funding program “Open Access Publication Funding”. Open Access funding enabled and organized by Projekt DEAL.

COMPETING INTERESTS

The authors declare no competing interests.

ETHICS APPROVAL

All experiments involving human materials adhered to the recommendations of the World Medical Association, as stated in the revised Declaration of Helsinki from Seoul 2008. The proposed project has been approved by the Ethics Commission of the University of Bonn (Lfd. Nr. 031/19). In addition, the applicants abided to all regulations specified in German Data Protection Law. Informed consent was obtained from all subjects involved in the study.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41431-022-01216-5>.

Correspondence and requests for materials should be addressed to Heiko Reutter or Alina C. Hilger.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022

3. Danksagung

Ich bedanke mich ganz herzlich bei meinem Doktorvater Prof. Dr. med. Heiko Reutter, meiner Betreuerin Dr. med. Alina Hilger und meinem Betreuer Dr. med. Gabriel Dworschak für die stetige Förderung, Ermutigung und Expertise, ohne die diese Doktorarbeit nicht möglich gewesen wäre. Des Weiteren bedanke ich mich bei den Mitarbeiter*innen unserer Arbeitsgruppe für die schöne Zeit im Labor. Ich danke den Patient*innen und Familien, die an unseren Studien teilgenommen haben. Bei meinen Eltern, meinem Bruder und meinen Freunden bedanke ich mich für die unermüdliche Geduld, ihr Verständnis und ihre beständige Unterstützung.