

ARTICLE

Copy number variations in 375 patients with oesophageal atresia and/or tracheoesophageal fistula

Erwin Brosens^{*,1,2,11}, Florian Marsch^{3,11}, Elisabeth M de Jong^{1,2}, Hitisha P Zaveri^{4,5}, Alina C Hilger³, Vera Gisela Choinitzki³, Alice Hölscher⁶, Per Hoffmann^{3,7}, Stefan Herms^{3,7}, Thomas M Boemers⁶, Benno M Ure⁸, Martin Lacher⁸, Michael Ludwig⁹, Bert H Eussen¹, Robert M van der Helm¹, Hannie Douben¹, Diane Van Opstal¹, Rene MH Wijnen², H Berna Beverloo¹, Yolande van Bever¹, Alice S Brooks¹, Hanneke IJsselstijn², Daryl A Scott^{4,5}, Johannes Schumacher³, Dick Tibboel², Heiko Reutter^{3,10,11} and Annelies de Klein^{1,11}

Oesophageal atresia (OA) with or without tracheoesophageal fistula (TOF) are rare anatomical congenital malformations whose cause is unknown in over 90% of patients. A genetic background is suggested, and among the reported genetic defects are copy number variations (CNVs). We hypothesized that CNVs contribute to OA/TOF development. Quantifying their prevalence could aid in genetic diagnosis and clinical care strategies. Therefore, we profiled 375 patients in a combined Dutch, American and German cohort via genomic microarray and compared the CNV profiles with their unaffected parents and published control cohorts. We identified 167 rare CNVs containing genes (frequency < 0.0005 in our in-house cohort). Eight rare CNVs – in six patients – were *de novo*, including one CNV previously associated with oesophageal disease. (hg19 chr7:g.(143820444_143839360)_(159119486_159138663)del) 1.55% of isolated OA/TOF patients and 1.62% of patients with additional congenital anomalies had *de novo* CNVs. Furthermore, three (15q13.3, 16p13.3 and 22q11.2) susceptibility loci were identified based on their overlap with known OA/TOF-associated CNV syndromes and overlap with loci in published CNV association case–control studies in developmental delay. Our study suggests that CNVs contribute to OA/TOF development. In addition to the identified likely deleterious *de novo* CNVs, we detected 167 rare CNVs. Although not directly disease-causing, these CNVs might be of interest, as they can act as a modifier in a multiple hit model, or as the second hit in a recessive condition.

European Journal of Human Genetics (2016) 24, 1715–1723; doi:10.1038/ejhg.2016.86; published online 20 July 2016

INTRODUCTION

Oesophageal atresia (OA) with or without tracheoesophageal fistula (TOF) are anatomical congenital malformations believed to be caused by multiple genetic and environmental factors.¹ With a prevalence of two to three in 10 000 live births, OA/TOF is a rare foregut-related anomaly.² Around 50% of affected individuals present with additional congenital anatomical malformations.³ Often – but not exclusively – these belong to the VATER/VACTERL association spectrum of vertebral defects (V), anorectal malformations (A), cardiac defects (C), TOF with or without OA (TE), renal anomalies (R) and radial limb defects (L).^{4,5}

A confirmed genetic syndrome or a chromosomal anomaly – including aneuploidies as trisomy 13, 18 and 21 – can be identified in 6–10% of patients,⁶ and there is a strong suspicion that genetic factors are involved in the remainder. A genetic background is further suggested by reports of families with multiple affected individuals, higher concordance rates in monozygotic twins compared with dizygotic twins,⁷ higher recurrence risk for siblings and children of affected individuals and OA/TOF as a component feature in numerous known chromosomal aberrations and monogenic syndromes.⁸ Reports describing disease-causing copy number

variations (CNVs) in patients with OA/TOF are rare.^{9,10} In addition to their well-established role in the development of congenital anatomical malformations in general,¹¹ CNVs contribute to disease aetiology in several genetic syndromes. These include those having OA/TOF as part of their phenotypic spectrum such as Feingold syndrome,¹² 22q11 deletion syndrome,¹³ CHARGE syndrome¹⁴ and mandibulofacial dysostosis.¹⁵ Furthermore, *de novo* disease-causing CNVs have been described in patients with non-syndromic OA/TOF and the VACTERL association.¹⁶

To determine the contribution of CNVs in OA/TOF aetiology, we profiled 375 Dutch, German and American OA/TOF patients in a comprehensive multiplatform array. We suggest that genomic *de novo* and rare overlapping CNVs contribute to isolated and non-isolated OA/TOF. These CNVs would harbour one or more disease-related genes or phenotype-modifying factors. We describe the variation detected in our large cohort. This study enabled us to identify several rare overlapping CNVs and nonoverlapping *de novo* CNVs, which potentially provide new insights into the biological pathways and disease mechanisms involved in the development of OA/TOF.

¹Department of Clinical Genetics, Erasmus Medical Centre – Sophia Children's Hospital, Rotterdam, The Netherlands; ²Department of Paediatric Surgery, Erasmus Medical Centre – Sophia Children's Hospital, Rotterdam, The Netherlands; ³Institute of Human Genetics, University of Bonn, Bonn, Germany; ⁴Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX, USA; ⁵Department of Molecular Physiology and Biophysics, Baylor College of Medicine, Houston, TX, USA; ⁶Department of Paediatric Surgery and Urology, Children's Hospital of Cologne, Cologne, Germany; ⁷Human Genomics Research Group, Department of Biomedicine, University Hospital and University of Basel, Basel, Switzerland; ⁸Center of Paediatric Surgery Hannover, Hannover Medical School and Bult Children's Hospital, Hannover, Germany; ⁹Department of Clinical Chemistry and Clinical Pharmacology, University of Bonn, Bonn, Germany; ¹⁰Department of Neonatology and Paediatric Intensive Care, University of Bonn, Bonn, Germany
*Correspondence: Dr E Brosens, Department of Clinical Genetics or Department of Paediatric Surgery, Erasmus Medical Centre – Sophia Children's Hospital, Room Ee971, PO Box 2040, 3000 CA Rotterdam, The Netherlands. Tel: +31 10 70 37643; Fax: +31 10 70 44736; E-mail: e.brosens@erasmusmc.nl

¹¹These authors contributed equally to this work.

Received 25 January 2016; revised 8 June 2016; accepted 14 June 2016; published online 20 July 2016

METHODS

Study design

We assessed the CNVs according to the consensus statement for chromosomal microarray analysis described by Miller *et al.*¹⁷ Our study design was based on the assumptions that CNVs are most likely to contribute to the abnormal phenotype in congenital anomalies if (I) a CNV is absent in large cohorts of unaffected individuals, (II) is absent in the unaffected parents of the affected individual and/or (III) is absent or has a population frequency below or comparable to the disease frequency, and (IV) if it targets relevant genes or noncoding RNAs. Recurrence of loci affected by *de novo* CNVs in single cases could indicate loci harbouring genes mutated or otherwise affected in larger disease cohorts. A detailed description of the study design is given in the Supplementary Methods.

Patient cohort

This study was approved by the institutional ethics committee of each participating centre, and was conducted in accordance with the principles of the Declaration of Helsinki. Patients with OA/TOF (isolated or non-isolated) were identified from the medical records. All patient records were reviewed by the treating physicians or geneticists of each participating centre. After retrieval of parental informed consent, blood was drawn from a total of 375 patients and their parents, comprising 239 patients from the Erasmus MC – Sophia, 28 from the Baylor College of Medicine and 108 from a German multicentre study regarding the genetic and environmental cause of OA/TOF ('The genetic risk for OA consortium (GREAT consortium)').

Microarray analysis

High-resolution analyses were performed using single-nucleotide polymorphism (SNP) microarrays (Illumina Inc., San Diego, CA, USA, and Affymetrix Inc., Santa Clara, CA, USA) and CGH oligonucleotide-based arrays (Agilent Inc., San Diego, CA, USA) using standard protocols. SNP data (log-R ratio, B-allele frequency) were visualized to identify potential CNVs via Biodiscovery Nexus CN7.5. (Biodiscovery Inc., Hawthorne, CA, USA) and the GenomeStudio genotyping module (v1.9.4, www.illumina.com). A detailed description of chip types, normalized output generation and analysis settings is provided in the Supplementary Methods. CNVs were – prior to validation studies – first filtered and prioritized based on size, probe content, quality, frequency in reference cohorts, gene content and frequency in our OA cohorts. All CNVs passing the filter criteria were evaluated manually in a modified version (ie, excluding BAC arrays and small InDels) of the database of genomic variation (<http://dgv.tcag.ca/dgv/app/home>), ISCA (<http://dbsearch.clinicalgenome.org/search/>), ClinGen (<https://www.clinicalgenome.org/data-sharing/clinvar/>) and DECIPHER (<http://decipher.sanger.ac.uk>). We classified CNVs to be rare if they were absent or present once in our in-house cohort of unaffected individuals ($n=3235$ individuals). We searched for overlap in large CNV cohorts of control individuals published by Cooper *et al.*,¹⁸ Coe *et al.*¹⁹ and Kaminsky *et al.*²⁰ We also evaluated the CNVs significantly different in these studies between patients and controls. To confirm the putative *de novo* and putative deleterious CNVs, patient and parental DNAs were tested with either additional SNP array, real-time quantitative PCR, fluorescence *in situ* hybridization (FISH) and/or multiplex amplicon quantification (MAQ; Multiplicon N.V., Gent, Belgium). A detailed description of these methods is given in the Supplementary Methods. All rare CNVs are listed in Supplementary Table 2 and are deposited in the ClinVar database (<http://www.ncbi.nlm.nih.gov/clinvar/>) using the submission name 'CNV study in EA/TEF' and using the exact identifiers as described in this manuscript.

RESULTS

Patient cohort

In this study a total of 375 OA/TOF patients were screened for their respective CNV profile. Of these patients, 129 presented with OA/TOF as an isolated defect (34.4%). Of the non-isolated patients with OA/TOF, 142 met the aforementioned criteria for the clinical diagnosis of VACTERL (37.8%).

Microarray analysis

Screening the respective cohorts (see Figure 1) with high-resolution oligonucleotide and SNP microarrays led to the identification of 169 CNV (gene-rich – containing genes – ($n=167$) and gene-poor ($n=2$)). These will be addressed as rare CNVs in the remainder of the manuscript. Their size distributions are depicted in Figure 2, genomic locations, evaluation of presence in control databases and classifications are given in Supplementary Table 2. Almost all of the rare CNVs were widely distributed over the genome. However, our analysis yielded a total of 12 loci that were affected by a rare CNV more than once and were present in more than one patient (see Supplementary Table 1 for the regions and phenotypes of patients with rare CNVs and overlapping loci). Inheritance was determined using secondary technology as MAQ assay or qPCR in 17 out of 74 CNVs either suspected to be *de novo* CNVs after trio-analysis using microarray or based on suspected deleteriousness in single-patient microarray analysis (see Supplementary Figure 1).

Eight out of these 74 rare CNVs selected for further investigation (10.8 %) – in six patients (1.6%) – were confirmed to be *de novo* (see Table 1, and Figures 3 and 4 for examples). In addition, one locus harboured a 15q11 *de novo* CNV deletion (hg19 chr15:g.(?_19339852-20216728_?), common in the database of genomic variants (see Supplementary Figure 1). All but one *de novo* CNVs were non-recurrent and nonoverlapping in our cohort. For four patients, DNA of only one parent was available, thus preventing determination of inheritance of the rare CNV in the missing parent. Haplotype analysis of the locus could confirm that the haplotype present in the patient was not the haplotype of the available parent in three out of four CNVs. In Table 1 the phenotypes of patients with confirmed *de novo* CNVs detected in this study and in Table 2 the *de novo* CNVs described in literature are shown. Most *de novo* CNVs described here and in literature are non-recurrent, that is, there are no overlapping loci. The only recurrent affected *de novo* locus is 7q35q36 (see Figure 3). One *de novo* CNV (16p13.3 duplication, see Table 1) overlapped two inherited 16p13.3 duplications (see Table 3). We classified the rare CNVs as benign (45), uncertain – likely benign (106) and uncertain (7). Interestingly, we could classify nine CNVs as uncertain-likely pathogenic and two as pathogenic. These putative deleterious CNVs seen in 10 patients (2.6%) are depicted in Table 3. Two of these were confirmed to be *de novo*, four were inherited from parents without OA and for four CNVs the inheritance pattern is not known.

DISCUSSION

We hypothesized that both *de novo* and rare overlapping CNVs could predispose to – or modify the phenotype of – OA/TOF patients. These disease-associated CNVs should be below or in the same frequency range as OA/TOF disease prevalence. We identified 169 of these rare CNVs including eight *de novo* CNVs (nonoverlapping) and 12 loci with overlapping rare gene-rich CNVs. Six patients in our cohort had rare CNVs confirmed to be *de novo*. The distribution of these *de novo* CNVs is comparable between isolated and non-isolated OA/TOF patients: two patients with isolated OA/TOF had one *de novo* CNV each (0.53% of total patient cohort; 1.55% of patients with isolated OA/TOF). Two patients with non-isolated OA/TOF had one *de novo* CNV each and two had two *de novo* CNVs (1.06% in total cohort; 1.62% of non-isolated OA patients). All *de novo* CNVs were non-recurrent in our cohort. However, there is overlap with structural chromosomal anomalies previously described in OA/TOF.²⁹ For instance, the chromosomal anomaly described by Jackson *et al.*³⁰ (46,XX,-13,+der(18)t(13;18)(q12;p11.2) overlaps with the 13q12 deletion detected in patient SKZ_1662. Genes in the deleted region may

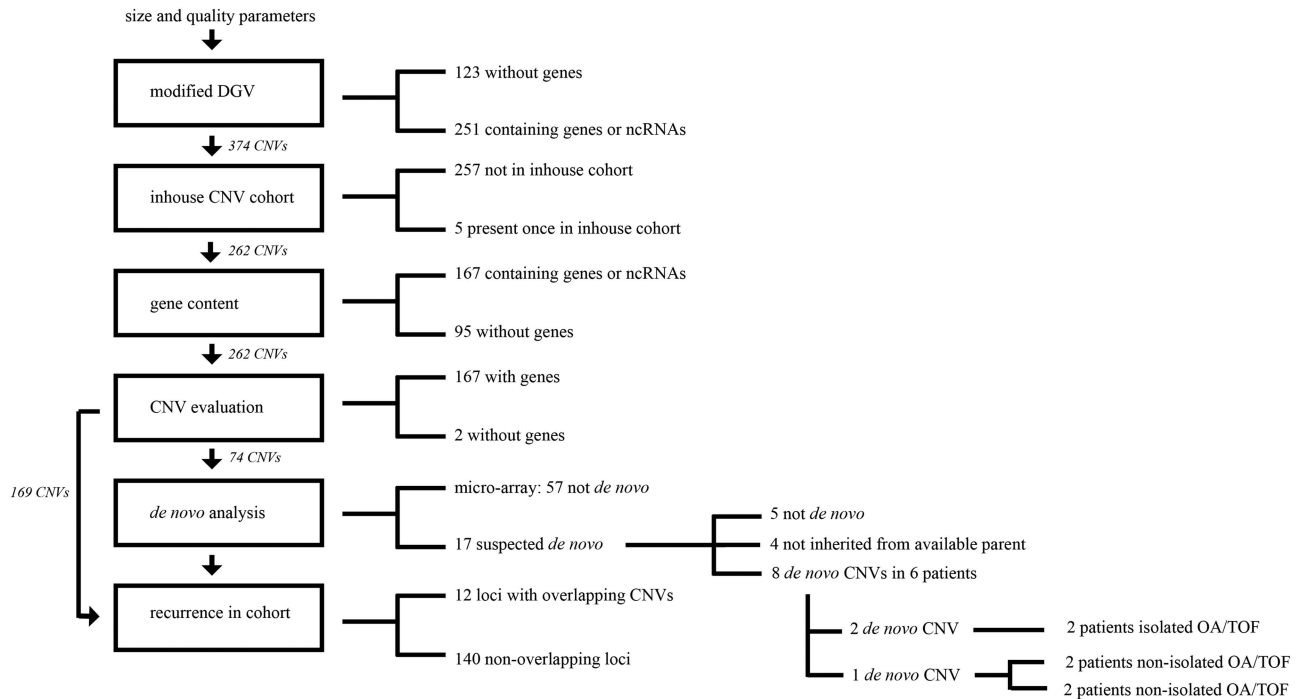


Figure 1 Filtering and prioritizing CNVs. After quality control and manual evaluation of CNVs, 374 CNVs larger than 30 kb, either absent or rare in the modified Database of Genomic Variants incorporated in the Nexus software, remained. Out of 374, 123 did not contain genes. In all, 257 were absent and 5 were present once in our in-house control database. These 262 CNVs were either gene-rich – containing genes – ($n=167$) or gene-poor ($n=95$). Two gene-poor CNVs were suspected of being *de novo* in microarray trio analysis. Eight out of 74 evaluated CNVs were *de novo*. Almost all of the rare CNVs (140) were widely distributed over the genome. However, our analysis yielded a total of 12 loci – containing 29 CNVs – which were affected by a rare CNV more than once and were present in more than one patient.

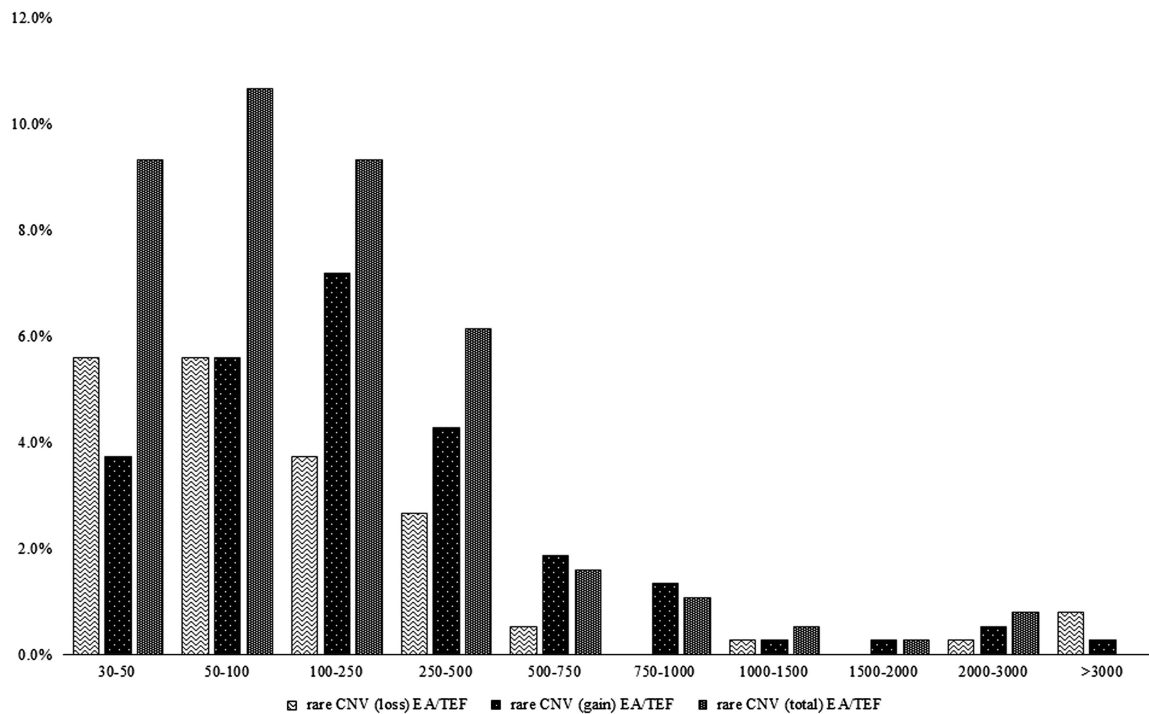


Figure 2 Size and type distribution of rare CNV. Total number of rare CNVs in the Erasmus MC – Sophia, Baylor College of Medicine and University of Bonn OA/TOF cohort (=375). Homozygous loss is counted as loss. Bins represent size ranges, for example, the 50–100 kb bin contains all CNVs within the size range of 50–100 kb.

Table 1 *De novo* CNVs in this cohort

Sample	Cytoband	CNV	(Candidate) genes	Phenotypes	Validation procedure	Classification
DE210SOUKBD100141	2q21.1	hg19 chr2:g.(131349899_131486169)_(131991166_132057166)dup	Multiple incl. <i>POTE1</i> , <i>GPR148</i> , <i>AMER3</i> , <i>ARHGEF4</i> , <i>FAM168B</i> , <i>PLEKH2</i>	OA/TOF	qPCR	Uncertain – likely benign
DE120SOUKBD100206	3p26.1	hg19 chr3:g.(8519329_8521841)_(8551649_8554052)dup	<i>LMCD1</i> , <i>LMCD1-AS1</i>	OA/TOF, tracheomalacia	qPCR	Uncertain
SKZ_2111	7q35q36.3	hg19 chr7:g.(143820444_143839360)_(159119486_159138663)del	Multiple incl. <i>SHH</i> , <i>MNX1</i> , <i>SLC4A2</i>	TOF, abnormally placed anus, hypospadias, ventricle septal defects	FISH	Pathogenic
SKZ_1810	8p22	hg19 chr8:g.(17613606_17625479)_(17613606_17844353)dup	<i>FGL1</i> , <i>LFIRE1</i> , <i>MTUS1</i> , <i>PGM1</i>	OA/TOF, aplastic left lung, anencephaly, cranio-achyrisis, kyphosis, ventricular septal defect, pulmonary stenosis, radial club hand, hypoplastic thumbs, horseshoe kidney, broad thorax	MAQ assay	Uncertain – likely benign
SKZ_1810	11p14.3	hg19 chr11:g.(21600567_21853276)_(22015682_22016670)del	-	OA/TOF, aplastic left lung, anencephaly, cranio-achyrisis, kyphosis, ventricular septal defect, pulmonary stenosis, radial club hand, hypoplastic thumbs, horseshoe kidney, broad thorax	MAQ assay	Uncertain – likely benign
SKZ_1662	13q12.11	hg19 chr13:g.(22685293_22688792)_(22978753_22981935)del	<i>AK054845</i> , <i>LINC00540</i>	OA/TOF, tracheal stenosis, abnormal sacrum and kidney abnormality	MAQ assay	Uncertain
SKZ_2111	16p13.11	hg19 chr16:g.(14985615_15156180)_(16289532_17000304)dup	Multiple	TOF, abnormally placed anus, hypospadias, ventricle septal defects	MAQ assay	Uncertain – likely pathogenic
SKZ_1307	21q22.2	hg19 chr21:g.(40099972_40100880)_(40152372_40154748)dup	<i>LINC00114</i>	OA/TOF, anal stenosis and long toes and fingers	MAQ assay	Uncertain – likely benign

None of the parents had oesophageal atresia or trachea-oesophageal fistula. Therefore, main focus was on *de novo* CNVs. Depicted are the eight confirmed (with either qPCR or MAQ assay) *de novo* CNV out of 74 evaluated rare CNVs from the Erasmus MC – Sophia, Baylor College of Medicine and University of Bonn cohorts. qPCR/FISH/MAQ assay validation results are in Supplementary Figure 1. Chromosome region according to build hg19.

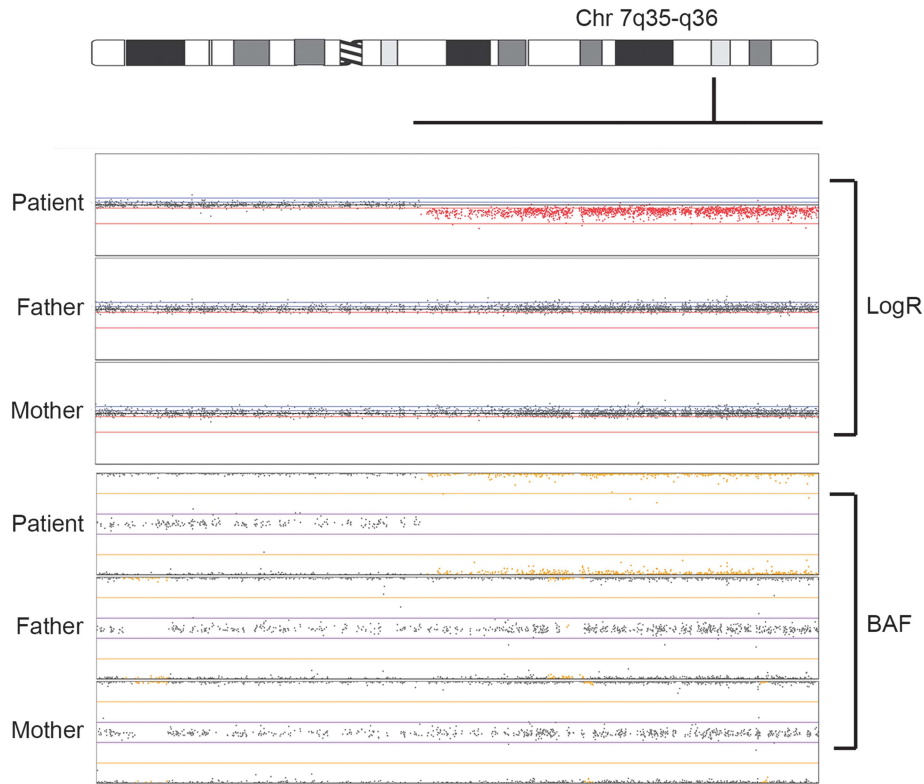


Figure 3 *De novo* deletion ranging from chromosomal band 7q35 to 7q36.3. Note the loss (red) in the patients logR track and the loss of heterozygosity (yellow) in the patients B-allege frequency (BAF) plot. qPCR/FISH/MAQ assay validation results in Supplementary Figure 1.

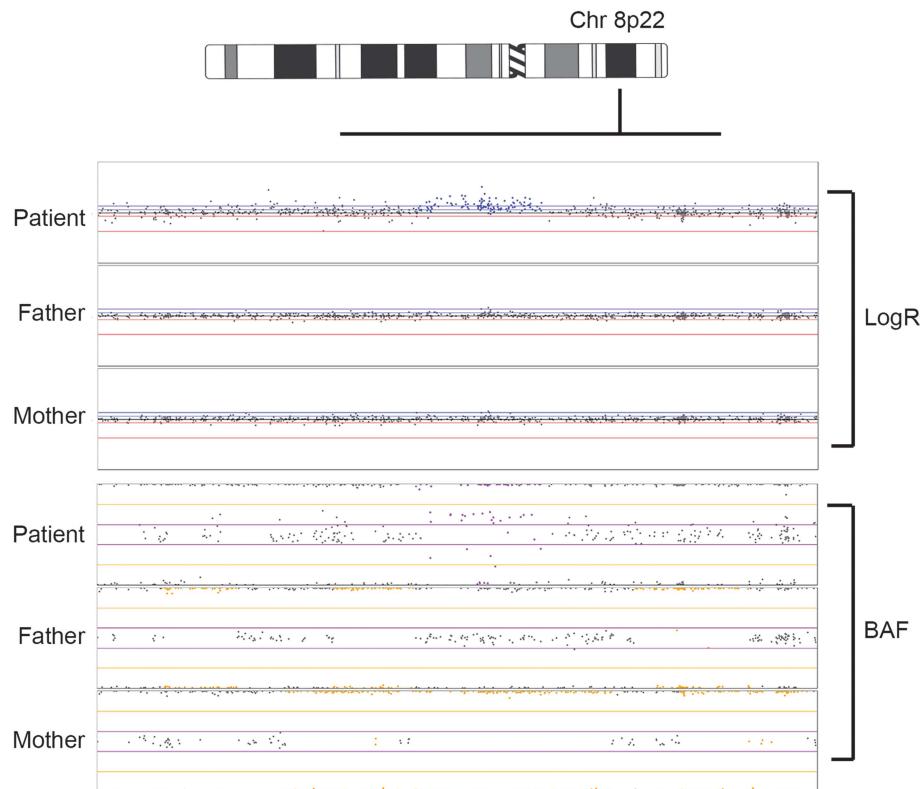


Figure 4 *De novo* duplication on chromosome 8p22. Note the gain (blue dots/arrow) in the patients' logR track and allelic imbalance (purple dots/arrow) in the patients BAF plot. qPCR/FISH/MAQ assay validation results in Supplementary Figure 1.

Table 2 De novo CNVs in OA/TOF patients described in literature

Cytoband	CNV	Minimum length (bp)	(Candidate) genes	Phenotypes	Reference
1q41	hg19 chr1:g.(?_217879151)_(218010441_?)dup	131 290	SPATA17	OA, butterfly vertebra, anal atresia without fistula, atrial and ventricular septal defects,	21
	hg19 chr8:g.(?_144940222)_(145060112_?)dup	119 890	PLEC-I, PARP10	bilateral dystopic kidneys, left-sided renal dysplasia	
2q33.2q35	hg19 chr2:g.(?_204,394,564)-(219,189,331_?)del	14 794 767	MAP2	OA, agenesis of umbilical artery, multiple brain abnormalities, tapered fingers, clinodactyly of the fifth fingers, deviated feet	22
3q28	hg19 chr3:g.(?_187913191)_(188468682_?)del	555 491	LPP	OA/TOF, tetralogy of Fallot, rib anomalies, hypospadias, small kidneys	23
7q35q36	hg19 chr7:g.(?_147683847)_(159088636_?)del	11 404 789	SHH, MNX1	Oesophageal stenosis, microcephaly, facial dysmorphism hypoplasia of the lower lumbar vertebrae, lumbosacral hyperlordosis, partial agenesis of sacrum, distended bladder, right ectopic kidney, developmental delay	24,25
11q13.1q13.2	hg19 chr11:g.(?_65508902)_(67473140_?)del	1 964 238	GSTP1	OA, facial dysmorphism, developmental delay	26
17q12	hg19 chr17:g.(?_34727386)_(36297053_?)dup	1 569 667	AATF, TADA2L, HNF1B	OA/TOF, anal atresia, sacral bone defect, cryptorchidism	27
20q13.33	hg19 chr20:g.(?_60238426)_(60895697_?)del	657 271	GTPBP5	OA/TOF, ventricular septal defects, hypospadias, large unilateral hydrocele, large left cystic hygroma	28

Depicted are the few published *de novo* CNVs in patients with nonsyndromic OA/TOF. Chromosome region according to build hg19.

Table 3 Putative deleterious rare CNVs in this cohort

Sample	Cytoband	CNV	Inheritance	Authors' classification	Phenotype
DE610SOUKBD100197	2q13	hg19 chr2:g.(111186302_111667198)_(113094793_113273657)dup	U	Uncertain – likely pathogenic	OA/TOF, sister as unilateral renal agenesis, both patient and sister are otherwise healthy
SKZ_1248	4q35.2	hg19 chr4:g.(187540292_187318091)_(187318091_187849681)dup	M	Uncertain – likely pathogenic	OA/TOF, pyloric stenosis; dysmorphisms, hearing loss
SKZ_1856	6p22.3	hg19 chr6:g.(20884837_20901311)_(20901267_21082258)del	NP	Uncertain – likely pathogenic	OA/TOF, septal defect, club foot
SKZ_1855	11p15.4	hg19 chr11:g.(4371631_4391231)_(51321119_5253127)dup	NP	Uncertain – likely pathogenic	OA/TOF
SKZ_0856	15q13.3	hg19 chr15:g.(32457092_32457092)_(32514341_32771537)del	P	Uncertain – likely pathogenic	OA/TOF, anal atresia, bifid/fused ribs, coarctation, abnormal arterial supply right lung, abnormal sacrum
SKZ_1150	16p13.11	hg19 chr16:g.(15539023_15545022)_(16282307_16291541)dup	P	Uncertain – likely pathogenic	OA/TOF+ atrioventricular septal defect
SKZ_1988	16p13.11	hg19 chr16:g.(15034035_15092778)_(15998820_16106095)dup	P	Uncertain – likely pathogenic	OA/TOF, anal anomalies
SKZ_1780	22q11.21	hg19 chr22:g.(18637139_18640300)_(20286099_20289862)dup	M	Uncertain – likely pathogenic	OA/TOF, anal atresia, ventricular septal defect
SKZ_0680	Xp22.2	hg19 chrX:g.(10299643_10302384)_(10637327_10638042)del	M	Pathogenic	Laryngo-tracheo-oesophageal-cleft, hypospadias, dysmorphisms, hypotonia, pyloric stenosis, opitz syndrome

Abbreviations: M, inheritance – maternal; NP, no parental DNA available; NZ, nullizygous; P, inheritance – paternal; U, inheritance – undetermined. Total number of putative deleterious CNVs in the Erasmus MC – Sophia, Baylor College of Medicine and University of Bonn OA/TOF cohort (= 375). Chromosome region according to build hg19. All CNVs were absent from our in-house control cohort.

contribute to the OA/TOF aetiology. Unfortunately, little is known about the genes within the region of overlap.

On the basis of the assumption that a CNV has a high likelihood of being pathogenic if it is not present in cohorts of unaffected individuals as well as absent in both unaffected parents, we could classify two out of eight *de novo* CNVs as (likely-) pathogenic, that is, those at the loci 16p13.11 and 7q35q36. No tracheoesophageal mouse phenotypes are described for any of the genes affected by *de novo* CNVs except for two genes in the 7q35q36 locus – *Shh* and *Slc4a2*. *Shh* knockout mice have numerous malformations including TOF, a small stomach, reduction of oesophageal tissue fused with the trachea, anal atresia and duodenal stenosis.³¹ *Slc4a2* knockout mice develop hyperkeratosis in oesophageal and stomach epithelium.³² As the 7q35q36 deletion is very large and contains many genes, other genes could also contribute to the abnormal phenotype seen in this patient. The remaining six *de novo* CNVs affected loci present in large CNV control cohorts (see supplementary table 2). The total number of patients with rare *de novo* CNVs is 6 out of 375 (1.6%). The *de novo* rate of 1.6% is slightly elevated compared with the *de novo* rate per genome/generation described by Itsara *et al.*^{33,34} They estimate a *de novo* rate of large CNVs to be 1.2%. However, these include more prevalent CNVs and not a selected subset. In other diseases, *de novo* CNVs have a much higher impact, for example, congenital diaphragmatic hernia³⁵ or intellectual disability.³⁶ The *de novo* CNVs seen in this study are nonoverlapping and some of them do not affect genes with clear association to the abnormalities seen in patients. Therefore, the significance of some of these *de novo* CNVs to the disease remains uncertain.

For instance, one *de novo* 13q11.2 deletion involves a female patient (SKZ_1662) born out of a twin pregnancy. Her twin died *in utero*. Unfortunately, there was neither patient material nor information regarding the observed congenital anomalies or zygosity status of this fetus available. The female index patient had OA/TOF, tracheal stenosis and a sacral abnormality, and her left kidney was abnormally positioned in the midline. Within the deleted 13q12.11 region one transcribed mRNA (*AK054845*) and one lncRNA (*LINC00540*) are located. No biological role or putative function has been described for these RNAs so far. However, one family – with congenital fibrosis of extraocular muscles – was reported having a translocation breakpoint (t(2;13)(q37.3;q12.11)) in this region. Mice *fgf9* knockouts showed a wide variety of abnormalities, including developmental problems of the skeletal, respiratory and the gastrointestinal systems.³⁷ The 13q12.11 *de novo* deletion observed in our patient is ~500 kb of *FGF9*, suggesting that a regulatory region of *FGF9* might be affected by the deletion. Female patient SKZ_1307 has a *de novo* duplication affecting a long noncoding RNA, *LINC00114*, on chromosome 21 (chr21:40100880-40154748) confirmed with MAQ assay. *LINC00114* is located between V-Ets Avian Erythroblastosis Virus E26 Oncogene Homolog (*ERG*) and V-Ets Avian Erythroblastosis Virus E26 Oncogene Homolog 2 (*ETS2*) within the Down's syndrome critical region.¹⁹ The girl has OA/TOF and anal stenosis as main additional features. She does not have distinct Down's syndrome facial features or mental retardation. The *ERG* and *ETS2* transcription factors might be regulated by *LINC00114*. Unfortunately, no mouse orthologue for this region exists.³⁸ *ERG* and *ETS2* are implicated as secondary hits – after an initial truncating *GATA1* mutation – in the development of neonatal transient myeloproliferative disease preceding myeloid leukaemia seen in Down's syndrome patients.³⁹ Patients with Down's syndrome have a higher prevalence of several gastrointestinal defects, including OA/TOF.⁴⁰ This is the first *de novo* duplication involving only one gene or long non-coding RNA in a patient with OA. Further investigation of the role of *LINC00114* in OA/TOF and Down's

syndrome patients with intestinal atresia is warranted. The identified *de novo* duplication on chromosome 3p26.1 in female patient DE12OSOUKBD100206 with OA and tracheomalacia comprises *LMCD1* encoding LIM and cysteine-rich domain protein 1, which acts as a transcriptional cofactor restricting the function of *GATA6*,⁴¹ a protein having an important role in endodermal differentiation.⁴² Moreover, *GATA6* expression has previously been reported to be elevated during the development and progression of Barrett's oesophagus in squamous epithelial cells.⁴³ Hence, the present finding of a *de novo* duplication comprising *LMCD1* in a patient with OA/TOF is suggestive of its pathogenic involvement in the development of OA/TOF. The importance and biological impact of the other *de novo* deletions/duplications is uncertain.

Of note, one *de novo* loss – a common polymorphism – was detected: hg19 chr15:g.(?_19339852)_(20216728_?)del. This CNV was detected during visual inspection of patient and parental SNP arrays for inheritance of other CNVs. This 15q11.2 polymorphism overlapped with a previously described genetic loss implicated in patients with congenital anatomical malformations, including OA/TOF.⁴⁴ This region is deleted in three more OA patients in our cohort.⁸ However, its high frequency in unaffected individuals and repetitive nature (eg, many LINE, SINE and other repetitive elements) hampers interpretation and classification of this CNV.

Overlapping rare CNVs

Rare CNVs are proposed to arise after replication errors¹¹ and have such a low population frequency that either they have arisen recently and have no biological meaning or are somehow detrimental and are virtually extinct from the population. Interpretation of these CNVs is difficult. For instance, they can be ancestry-specific.⁴⁵ Inheritance of a single CNV from a healthy parent is generally a characteristic of a benign CNV. However, absence of distinct abnormalities in parents carrying the same rare CNV could, for instance, be explained by a subclinical phenotype in these parents, variable gene expressivity, incomplete penetrance, skewed X-inactivation and/or mutations elsewhere in the genome.¹¹ Reduced penetrance or variable expressivity of CNVs has been described in patients with OA/TOF. For instance, Faguer *et al.*⁴⁶ described differences in expression of a microduplication in patients with the same microduplication, a father with bilateral vesicoureteric reflux and renal hypodysplasia and his child with left multicystic dysplastic kidney with megaureter, vesicoureteric reflux, bladder diverticulae and OA/TOF. Both patients have the same duplication on chromosomal locus 17q12, which includes *HNF1B*, a gene mutated in one-fifth of patients with dysplastic kidneys.⁴⁶

The best way to see whether a CNV is associated to a disease is to do a formal burden test.⁴⁷ We are not able to do this test because of the limited number of patients in a rare disease, and because of technical limitations (use of different array chips). More details are given in the supplementary discussion. However, we can look for overlap with CNVs described previously in CNV burden studies and inspect whether OA/TOF has been described in patients with such a CNV. Therefore, we used the CNV burden studies published by Cooper *et al.*,¹⁸ Coe *et al.*¹⁹ and Kaminsky *et al.*²⁰ as a proxy (developmental delay *vs* controls) after filtering all common CNVs. Here, they did use sufficient numbers of patients and controls and find an enrichment of a small number of loci in this heterogeneous patient population of developmental delay and/or congenital anomalies. Only the 16p13.3 duplication enriched in patients in these studies was recurrent in our cohort. The largest of the three duplications – seen in patient SKZ_2111 – was *de novo*. The two other paternally inherited 16p13.3 duplications were present in patients SKZ_1988 and

SKZ_1150. Duplications of this region between the *NOMO1* and *XYLT1* gene have been described previously in patients with various phenotypical anomalies, including the OA/TOF-associated congenital anomalies and anal and cardiac malformations.⁴⁸ None of the other overlapping rare CNVs found in our cohort (see supplementary table 1) was enriched in the developmental delay study.

Non-recurrent CNVs seen in our cohort did overlap enriched CNVs in these burden studies or with CNVs published in patient databases. For instance, the 15q13.3 deletion seen in male patient SKZ_0856 overlaps with a known deleterious CNV¹⁸ seen in patients with a highly variable phenotype, which include mild to moderate intellectual disability and variable dysmorphic features.⁴⁹ Other CNVs with overlap in our study are the gain involving *FAT1* on 4q35.2 in patient SKZ_1248, the 6p22 deletion in patient SKZ_1856,¹⁸ the 2q13 duplication seen in patient DE61OSOUKBD100197¹⁹ and 22q11 gain¹⁸ seen in female patient SKZ_1780. Interestingly, two additional published EA/TEF patients have a 22q11 duplication overlapping the one seen in patient SKZ_1780. The DECIPHER database contains an inherited gain (chr22:19095778-19928090) described in patient 3771, with TOF, upper respiratory tract abnormality, coloboma, hearing impairment, horseshoe kidney and a right aortic arch with mirror image branching. The second is a paternally inherited duplication in a patient with OA/TOF and ventricular septal defect.⁵⁰

Rare CNVs could be determinants in secondary phenotypical anomalies and/or serve as a second 'hit' tilting the balance from normal to abnormal development. Duplications might be rescue mechanisms in which a normal copy is duplicated to balance out a copy affected by a mutation, resulting in increased gene expression or deletions might worsen an otherwise less severe condition. OA/TOF is a variable feature in several single-gene disorders. Perhaps the presence of these disorders is higher than currently diagnosed. Recognizing the phenotypical spectra might be hampered by uncharacteristic phenotypical features in patients carrying both a modifying rare CNV and a gene mutation. It might be worthwhile to screen large OA/TOF patient cohorts retrospectively for mutations in known disease genes. Unfortunately, owing to the large number of genes and non-recurrence of *de novo* CNVs, it is not feasible to establish their contribution to OA/TOF disease aetiology. Moreover, the lack of availability of OA/TOF patient samples and heterogeneity of the rare CNVs hamper formal burden analysis to prove association. However, the *de novo* nature of CNVs in patients and absence of overlapping CNVs in a large control cohorts is interesting. Perhaps, future CNV profiling or sequencing studies will detect deleterious variation in overlapping genes, paving the way for further single-gene-based functional studies.

Concluding remarks

We hypothesized that *de novo* and overlapping rare recurrent CNVs could contribute to the disturbed development of the oesophagus. Quantifying CNV prevalence and identity could aid in genetic diagnosis and clinical care selection. We found several *de novo* and rare overlapping CNVs. Our screening indicated that the prevalence of *de novo* CNVs on OA/TOF patient population is 1.6%. On the basis of their function, overlap with loci in published case-control studies, known CNV syndromes and foregut phenotypes in animal models, we suggest *SHH* and *SLC4A2* as contributing factors in a contiguous gene deletion to OA/TOF disease aetiology, and 15q13.3, 16p13.3 and 22q11.2 as candidate susceptibility loci. With aneuploidy and structural chromosomal anomalies (~4%) and single base pair mutations (~6%) now CNVs (~1–2%) total the genetic contribution of OA/TOF disease aetiology to ~11–12%. Mutation screening using candidate

gene approaches, whole-exome or whole-genome sequencing as well as sequencing large patient-parent cohorts – both prospectively and retrospectively – will likely reveal known and new pathogenic DNA variations, increasing the contribution of genetics and our knowledge of OA/TOF disease aetiology.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

We thank all families and patients for their cooperation, and the self-help organizations from Germany 'KEKS e.V.' and the Netherlands 'VOKS' for their assistance with recruitment. We thank David Alexander for editorial support. EB was funded by the Sophia Foundations for Scientific Research, projects SSWO- 493 and SWOO13-09. HR and JS are funded by a grant of the Else Kröner-Fresenius-Stiftung (EKFS; funding code 2014_A14). A full list of centres that contributed to the generation of the Decipher community data is available from <http://decipher.sanger.ac.uk> and via email from decipher@sanger.ac.uk. This project was funded by the Wellcome Trust.⁷

- 1 Felix JF, de Jong EM, Torfs CP, de Klein A, Rottier RJ, Tibboel D: Genetic and environmental factors in the etiology of oesophageal atresia and/or tracheoesophageal fistula: an overview of the current concepts. *Birth Defects Res* 2009; **85**: 747–754.
- 2 Pedersen RN, Calzolari E, Husby S, Garne E: Oesophageal atresia: prevalence, prenatal diagnosis and associated anomalies in 23 European regions. *Arch Dis Child* 2012; **97**: 227–232.
- 3 Genevieve D, de Pontual L, Amiel J, Sarnacki S, Lyonnet S: An overview of isolated and syndromic oesophageal atresia. *Clin Genet* 2007; **71**: 392–399.
- 4 Quan L, Smith DW: The VATER association. Vertebral defects, anal atresia, T-E fistula with oesophageal atresia, radial and renal dysplasia: a spectrum of associated defects. *J Pediatr* 1973; **82**: 104–107.
- 5 Temtamy SA, Miller JD: Extending the scope of the VATER association: definition of the VATER syndrome. *J Pediatr* 1974; **85**: 345–349.
- 6 Genevieve D, de Pontual L, Amiel J, Lyonnet S: Genetic factors in isolated and syndromic oesophageal atresia. *J Pediatr Gastroenterol Nutr* 2011; **52** (Suppl 1): S6–S8.
- 7 Schulz AC, Bartels E, Stressig R *et al*: Nine new twin pairs with oesophageal atresia: a review of the literature and performance of a twin study of the disorder. *Birth Defects Res* 2012; **94**: 182–186.
- 8 Brosens E, Ploeg M, van Bever Y *et al*: Clinical and etiological heterogeneity in patients with tracheo-oesophageal malformations and associated anomalies. *Eur J Med Genet* 2014; **57**: 440–452.
- 9 Reutter H, Ludwig M: VATER/VACTERL association: evidence for the role of genetic factors. *Mol Syndromol* 2013; **4**: 16–19.
- 10 Bednarczyk D, Sasiadek MM, Smigiel R: Chromosome aberrations and gene mutations in patients with oesophageal atresia. *J Pediatr Gastroenterol Nutr* 2013; **57**: 688–693.
- 11 Stankiewicz P, Lupski JR: Structural variation in the human genome and its role in disease. *Annu Rev Med* 2010; **61**: 437–455.
- 12 Cognet M, Nougayre A, Malan V *et al*: Dissection of the MYCN locus in Feingold syndrome and isolated oesophageal atresia. *Eur J Hum Genet* 2011; **19**: 602–606.
- 13 Digilio MC, McDonald-McGinn DM, Heike C *et al*: Three patients with oculo-auriculo-vertebral spectrum and microdeletion 22q11.2. *Am J Med Genet A* 2009; **149A**: 2860–2864.
- 14 Corsten-Janssen N, Saitta SC, Hoefsloot LH *et al*: More clinical overlap between 22q11.2 deletion syndrome and CHARGE syndrome than often anticipated. *Mol Syndromol* 2013; **4**: 235–245.
- 15 Gordon CT, Petit F, Oufadem M *et al*: EFTUD2 haploinsufficiency leads to syndromic oesophageal atresia. *J Med Genet* 2012; **49**: 737–746.
- 16 Brosens E, Eussen H, van Bever Y *et al*: VACTERL association etiology: the impact of *de novo* and rare copy number variations. *Mol Syndromol* 2013; **4**: 20–26.
- 17 Miller DT, Adam MP, Aradhya S *et al*: Consensus statement: chromosomal microarray is a first-tier clinical diagnostic test for individuals with developmental disabilities or congenital anomalies. *Am J Hum Genet* 2010; **86**: 749–764.
- 18 Cooper GM, Coe BP, Girirajan S *et al*: A copy number variation morbidity map of developmental delay. *Nat Genet* 2011; **43**: 838–846.
- 19 Coe BP, Witherspoon K, Rosenfeld JA-O *et al*: Refining analyses of copy number variation identifies specific genes associated with developmental delay. *Nat Genet* 2014; **46**: 1063–1071.
- 20 Kaminsky EB, Kaul V, Paschall J *et al*: An evidence-based approach to establish the functional and clinical significance of copy number variants in intellectual and developmental disabilities. *Genet Med* 2011; **13**: 777–784.
- 21 Hilger A, Schramm C, Pennimpede T *et al*: *De novo* microduplications at 1q41, 2q37.3, and 8q24.3 in patients with VATER/VACTERL association. *Eur J Hum Genet* 2013; **21**: 1377–1382.

- 22 van Binsbergen E, Ellis RJ, Abdelmalik N *et al*: A fetus with *de novo* 2q33.2q35 deletion including MAP2 with brain anomalies, esophageal atresia, and laryngeal stenosis. *Am J Med Genet* 2014; **164a**: 194–198.
- 23 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* 2010; **152a**: 2919–2923.
- 24 Zen PR, Riegel M, Rosa RF *et al*: Esophageal stenosis in a child presenting a *de novo* 7q terminal deletion. *Eur J Med Genet* 2010; **53**: 333–336.
- 25 Speleman F, Van Roy N, Wiegant J *et al*: Detection of subtle reciprocal translocations by fluorescence *in situ* hybridization. *Clin Genet* 1992; **41**: 169–174.
- 26 Ferreira de Almeida T, Bertola DR: Microdeletion 11q13.1.q13.2 in a patient presenting with developmental delay, facial dysmorphism, and esophageal atresia: possible role of the GSTP1 gene in esophagus malformation. *Birth Defects Res* 2013; **97**: 463–466.
- 27 Smigielski R, Marcellis C, Patkowski D *et al*: Esophageal atresia with tracheoesophageal fistula and anal atresia in a patient with a *de novo* microduplication in 17q12. *Eur J Med Genet* 2014; **57**: 40–43.
- 28 Solomon BD, Pineda-Alvarez DE, Hadley DW *et al*: *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. *Birth Defects Res* 2011; **91**: 862–865.
- 29 Felix JF, Tibboel D, de Klein A: Chromosomal anomalies in the aetiology of oesophageal atresia and tracheo-oesophageal fistula. *Eur J Med Genet* 2007; **50**: 163–175.
- 30 Jackson J, Robson L, Meagher S, Watson G, Smith A: How accurate does rapid fetal karyotyping need to be? Case of unbalanced t(13;18). *Prenat Diagn* 1993; **13**: 767–770.
- 31 Miller LA, Wert SE, Clark JC, Xu Y, Perl AK, Whitsett JA: Role of Sonic hedgehog in patterning of tracheal-bronchial cartilage and the peripheral lung. *Dev Dyn* 2004; **231**: 57–71.
- 32 Gawenis LR, Ledoussal C, Judd LM *et al*: Mice with a targeted disruption of the AE2 Cl⁻/HCO₃⁻ exchanger are achlorhydric. *J Biol Chem* 2004; **279**: 30531–30539.
- 33 Itsara A, Wu H, Smith JD *et al*: *De novo* rates and selection of large copy number variation. *Genome Res* 2010; **20**: 1469–1481.
- 34 Itsara A, Cooper GM, Baker C *et al*: Population analysis of large copy number variants and hotspots of human genetic disease. *Am J Hum Genet* 2009; **84**: 148–161.
- 35 Wat MJ, Veenma D, Hogue J *et al*: Genomic alterations that contribute to the development of isolated and non-isolated congenital diaphragmatic hernia. *J Med Genet* 2011; **48**: 299–307.
- 36 Hehir-Kwa JY, Rodriguez-Santiago B, Vissers LE *et al*: *De novo* copy number variants associated with intellectual disability have a paternal origin and age bias. *J Med Genet* 2011; **48**: 776–778.
- 37 Harada M, Murakami H, Okawa A *et al*: FGF9 monomer-dimer equilibrium regulates extracellular matrix affinity and tissue diffusion. *Nat Genet* 2009; **41**: 289–298.
- 38 Owczarek CM, Portbury KJ, Hardy MP *et al*: Detailed mapping of the ERG-ETS2 interval of human chromosome 21 and comparison with the region of conserved synteny on mouse chromosome 16. *Gene* 2004; **324**: 65–77.
- 39 Gruber TA, Downing JR: The biology of pediatric acute megakaryoblastic leukemia. *Blood* 2015; **126**: 943–949.
- 40 Freeman SB, Torfs CP, Romitti PA *et al*: Congenital gastrointestinal defects in Down syndrome: a report from the Atlanta and National Down Syndrome Projects. *Clin Genet* 2009; **75**: 180–184.
- 41 Rath N, Wang Z, Lu MM, Morrissey EE: LMCD1/Dyxin is a novel transcriptional cofactor that restricts GATA6 function by inhibiting DNA binding. *Mol Cell Biol* 2005; **25**: 8864–8873.
- 42 Molkenin JD: The zinc finger-containing transcription factors GATA-4, -5, and -6. Ubiquitously expressed regulators of tissue-specific gene expression. *J Biol Chem* 2000; **275**: 38949–38952.
- 43 Pavlov K, Honing J, Meijer C *et al*: GATA6 expression in Barrett's oesophagus and oesophageal adenocarcinoma. *Dig Liver Dis* 2015; **47**: 73–80.
- 44 Wong D, Johnson SM, Young D, Iwamoto L, Sood S, Slavin TP: Expanding the BP1-BP2 15q11.2 microdeletion phenotype: tracheoesophageal fistula and congenital cataracts. *Case Rep Genet* 2013; **2013**: 801094.
- 45 Oei L, Hsu YH, Stykarsdottir U *et al*: A genome-wide copy number association study of osteoporotic fractures points to the 6p25.1 locus. *J Med Genet* 2013; **51**: 122–131.
- 46 Faguer S, Chassaing N, Bandin F *et al*: A 17q12 chromosomal duplication associated with renal disease and esophageal atresia. *Eur J Med Genet* 2011; **54**: e437–e440.
- 47 Barnes C, Plagnol V, Fitzgerald T *et al*: A robust statistical method for case-control association testing with copy number variation. *Nat Genet* 2008; **40**: 1245–1252.
- 48 Nagamani SC, Erez A, Bader P *et al*: Phenotypic manifestations of copy number variation in chromosome 16p13.11. *Eur J Hum Genet* 2011; **19**: 280–286.
- 49 van Bon BW, Mefford HC, Fau - Menten B, Menten B, Fau - Koolen DA *et al*: Further delineation of the 15q13 microdeletion and duplication syndromes: a clinical spectrum varying from non-pathogenic to a severe outcome. *J Med Genet* 2009; **46**: 511–523.
- 50 Puvabanditsin S, Garrow E, February M, Yen E, Mehta R: Esophageal atresia with recurrent tracheoesophageal fistulas and microduplication 22q11.23. *Genetic Counsel (Geneva, Switzerland)* 2015; **26**: 313–320.

Supplementary Information accompanies this paper on European Journal of Human Genetics website (<http://www.nature.com/ejhg>)