COLD SPRING HARBOR
Molecular Case Studies

RESEARCH REPORT

A t(5;16) translocation is the likely driver of a syndrome with ambiguous genitalia, facial dysmorphism, intellectual disability, and speech delay

Ayşegül Ozantürk,[1,2,3] Erica E. Davis,[3] Aniko Sabo,[4] Marjan M. Weiss,[5] Donna Muzny,[4] Shannon Dugan-Perez,[4] Erik A. Sistermans,[5] Richard A. Gibbs,[4] Köksal R. Özgül,[1] Dilek Yalnızoglu,[6] Esra Serdaroglu,[6] Ali Dursun,[1] and Nicholas Katsanis[3]

[1]Department of Pediatrics, Metabolism Unit, Hacettepe University, Ankara 06410, Turkey; [2]Department of Biology, Molecular Biology Section, Hacettepe University, Ankara 06800, Turkey; [3]Center for Human Disease Modeling, Duke University, Durham, North Carolina 27701, USA; [4]Human Genome Sequencing Center, Baylor College of Medicine, Houston, Texas 77030, USA; [5]Department of Clinical Genetics, VU University Medical Center (Amsterdam), NL-1081 HV Amsterdam, The Netherlands; [6]Department of Pediatrics, Pediatric Neurology Unit, Hacettepe University, Ankara 06410, Turkey

Abstract Genetic studies grounded on monogenic paradigms have accelerated both gene discovery and molecular diagnosis. At the same time, complex genomic rearrangements are also appreciated as potent drivers of disease pathology. Here, we report two male siblings with a dysmorphic face, ambiguous genitalia, intellectual disability, and speech delay. Through quad-based whole-exome sequencing and concomitant molecular cytogenetic testing, we identified two copy-number variants (CNVs) in both affected individuals likely arising from a balanced translocation: a 13.5-Mb duplication on Chromosome 16 (16q23.1 → 16qter) and a 7.7-Mb deletion on Chromosome 5 (5p15.31 → 5pter), as well as a hemizygous missense variant in *CXorf36* (also known as *DIA1R*). The 5p terminal deletion has been associated previously with speech delay, whereas craniofacial dysmorphia and genital/urinary anomalies have been reported in patients with a terminal duplication of 16q. However, dosage changes in either genomic region alone could not account for the overall clinical presentation in our family; functional testing of *CXorf36* in zebrafish did not induce defects in neurogenesis or the craniofacial skeleton. Notably, literature and database analysis revealed a similar dosage disruption in two siblings with extensive phenotypic overlap with our patients. Taken together, our data suggest that dosage perturbation of genes within the two chromosomal regions likely drives the syndromic manifestations of our patients and highlight how multiple genetic lesions can contribute to complex clinical pathologies.

[Supplemental material is available for this article.]

Corresponding authors:
adursun@hacettepe.edu.tr;
katsanis@cellbio.duke.edu

Ontology terms: cat cry; central hypotonia; clubbing of toes; delayed gross motor development; down-sloping shoulders; enlarged proximal interphalangeal joints; high, narrow palate; intellectual disability, profound; micropenis; moderately short stature; penile hypospadias; pes planus; prominent forehead; thin upper lip vermilion; wide nasal bridge

Published by Cold Spring Harbor Laboratory Press

doi: 10.1101/mcs.a000703

INTRODUCTION

Whole-exome sequencing (WES) has become a commonly used tool for both gene discovery in the research setting and accelerated diagnostics in the clinical setting. Reports from ever-increasing patient cohorts have indicated that WES can deliver a molecular diagnosis ~25% of the time (Yang et al. 2013; Beaulieu et al. 2014). In parallel, the combinatorial

application of oligo-based comparative genomic hybridization arrays and computational analyses that infer chromosomal losses/gains from nonuniform read depth have also highlighted an important contribution of copy-number variants (CNVs) to rare genetic disorders (Solinas-Toldo et al. 1997; Snijders et al. 2001; Schaaf et al. 2011).

Despite these advances, a significant fraction of patients tested by these methods yields no overt candidates, even in cases in which the inheritance of the phenotype within a family suggests a strong underlying genetic component. There are many reasons for this, including the persistent difficulties in interpreting rare variation, the low quality or coverage of some regions of the genome, the poor detection of some CNVs, and the contribution of mutations in noncoding regulatory elements. In addition, examples have emerged in which a single-gene–single-syndrome paradigm is not sufficient to explain the inheritance of clinical phenotypes. In some cases, the clinical presentation represents two overlapping disorders, thus confusing locus assignment (Ng et al. 2010; Roach et al. 2010; Yang et al. 2013). In other cases, alleles can interact with each other at discrete loci to modulate penetrance and/or expressivity. These can represent genetic interactions of either single-nucleotide variants (SNVs) (Davis and Katsanis 2012), SNVs with CNVs (Lemmers et al. 2012; Lindstrand et al. 2014), or, in principle, CNVs with CNVs (Girirajan et al. 2012).

Here, we report the combinatorial results of quad-based WES and molecular cytogenetic analysis of Turkish male siblings with a constellation of neurocognitive, craniofacial, and gonadal malformations that do not conform readily to an established syndrome or known molecular lesion. The aggregate analysis for SNVs and CNVs identified two dosage perturbations in both affected individuals—a terminal duplication on Chromosome 16q and a terminal deletion on Chromosome 5p—as well as a maternally inherited missense variant (p.Ser105Pro) in *CXorf36*. Functional studies in zebrafish embryos indicated that perturbation of *CXorf36* is unlikely to contribute to the patients' dysmorphic features or cause neurodevelopmental abnormalities, whereas molecular cytogenetic analysis indicated that no gene was disrupted in either CNV boundary. In contrast, retrospective analysis of reported cases led to the identification of a family of Saudi Arabian descent who shared clinical features with our patients and had likewise overlapping 5p loss and 16q gain (Hellani et al. 2010). Some clinical features can potentially be explained by dosage defects on either chromosomal location. However, the remaining defects in our patients, which are shared by the previously published family, argue that concomitant haploinsufficiency on 5p and increased gene dosage across 16q constitute the most parsimonious driver hypothesis for this syndrome.

RESULTS

Clinical Characterization of a Syndromic Disorder in a Turkish Pedigree

We consulted for a nonconsanguineous Turkish family with two male siblings with the primary features of motor delay accompanied by intellectual disability and ambiguous genitalia (Table 1; Fig. 1A) at Hacettepe University Hospital. The parents reported no family history of inherited disease, except for a deceased female child who was affected with hypotonia of unknown etiology. We evaluated the oldest affected child (M-11-1496) when he was 11 yr, 4 mo old. Physical examination showed a weight of 26.5 kg (<5th percentile), height of 124 cm (<5th percentile), and head circumference of 51 cm. He had a weak cry (cat-like), stridor, and hypotonia at birth; his stridor recovered when he was 18 mo of age. He had delayed developmental milestones: He was able to hold his head up and sit by himself at 4 yr and he walked at 9.5 yr. First, we noted focal neurological deficits and stereotypic movements. He did not respond to his name nor make eye contact, and he was unable to speak. His gait is wide because of pes planus, and he displayed minimal coordination. Second, he has a micropenis, hypospadias, and undescended testes. Moreover, we noted distinct craniofacial and

Table 1. Phenotypic comparison between individuals with 16q23 dup and 5p15 del

	Present report		Hellani et al. 2010	
Phenotype	M-11-1496 16q23.1 → ter 5p15.31 → ter	M-11-1497 16q23.1 → ter 5p15.31 → ter	Patient 1 16q23.3→ter 5p15.32→ter	Patient 2 16q23.3→ter 5p15.32→ter
Hypotonia ***HP:0001319***	+	+	+	+
Congenital heart disease *HP:0001635*	–	–	+	+
Umbilical hernia *HP:0001537*	–	–	+	+
Pectus excavatum *HP:0000767*	–	–	+	+
Genital hypoplasia ***HP:0003241***	+	+	+	+
Short stature ***HP:0004322***	+	+	+	+
Hip rotation *HP:0001374*	–	–	+	+
Metatarsus adductus *HP:0001840*	–	–	+	+
Clinodactyly *HP:0030084*	–	–	+	+
Motor developmental delay ***HP:0001270***	+	+	+	+
Intellectual disability ***HP:0001249***	+	+	+	+
Brachycephaly *HP:0000248*	–	–	+	+
Prominent forehead ***HP:0011220***	+	+	+	+
Broad nasal bridge ***HP:0000431***	+	+	+	+
Low-set ears ***HP:0000369***	+	+	+	+
Arched palate ***HP:0000218***	+	+	+	+
Thin upper lip ***HP:0000177***	+	+	+	+
Esotropia *HP:0000565*	NR	NR	+	+
Speech delay ***HP:0000750***	+	+	+	+
Stridor *HP:0010307*	+	+	NR	NR
Scaphocephaly *HP:0000268*	+	+	NR	NR
Stereotypic movements *HP:000733*	+	+	NR	NR

Continued

Table 1. *Continued*

Phenotype	Present report		Hellani et al. 2010	
	M-11-1496 16q23.1 → ter 5p15.31 → ter	M-11-1497 16q23.1 → ter 5p15.31 → ter	Patient 1 16q23.3→ter 5p15.32→ter	Patient 2 16q23.3→ter 5p15.32→ter
Cat-like cry *HP:0200046*	+	–	–	–
Clubbing *HP:0001217*	+	+	NR	NR
Swollen interphalangeal joints *HP:0006253*	+	+	NR	NR
Distal lower limb amyotrophy *HP:0008944*	+	+	NR	NR
Narrow shoulders ***HP:0000774***	+	+	+	+
Pes planus *HP:0001763*	+	+	NR	NR

Human Phenotype Ontology (HPO) terms are shown in italics. Phenotypes in bold are common features between presented patients and reported patients in Hellani et al. 2010.
+, present; –, absent; NR, not recorded.

skeletal abnormalities. He had mild scaphocephaly and a distinct facial gestalt, with an elongated face, prominent forehead with a narrow diameter, high arched palate, and low-set ears (the family declined the use of photographs). Additionally, his shoulders are narrow; he has clubbing, thenar and hypothenar types of neural atrophy of his hands, swelling around the interphalangeal joints, and distal atrophy in his lower extremities. Multiple clinical laboratory tests were normal, including blood chemistry, lipid profile, creatinine kinase, routine urine tests, amino acid profile, and biotinidase activity. A hearing test and electromyography were both normal.

We observed multiple overlapping phenotypes between M-11-1496 and his younger brother (M-11-1497), who was 8.5 yr old at the time of evaluation. He was born prematurely at 32 wk, he was hypotonic, and his sucking reflex was immature. Additionally, he had stridor that resolved at 12 mo. He was able to hold his head up at 18 mo and walked at 3 yr. A physical examination showed a weight of 19 kg (<5th percentile), height of 114 cm (<5th percentile), and head circumference of 50 cm. His neurological findings included motor delay, hypotonia, stereotypic movements, and intellectual disability; he could speak three words after he had speech therapy for 2 yr. He had hypoactive deep tendon reflexes, flexor planar reflexes, and limited eye contact. Also similar to his brother, he displayed genital malformations: He had micropenis, hypospadias, and undescended testes. Further, he had craniofacial and skeletal abnormalities similar to those of his brother. All routine clinical laboratory tests (blood count, lipid profile, amino acid profile, and biotinidase activity) were normal except for mild anemia (hemoglobin [Hb], 10.6 g/dL; mean corpuscular volume [MCV], 79.6).

WES Reveals a Single X-Linked Candidate

To identify the molecular basis of the syndromic features of the affected male siblings, DNA samples from all four individuals in the family were collected; DNA was unavailable from the deceased sister for analysis. The entire quad was subjected to WES, whereas the two patients also underwent clinical aCGH (array comparative genomic hybridization) analysis, with the results of the two studies merged post hoc.

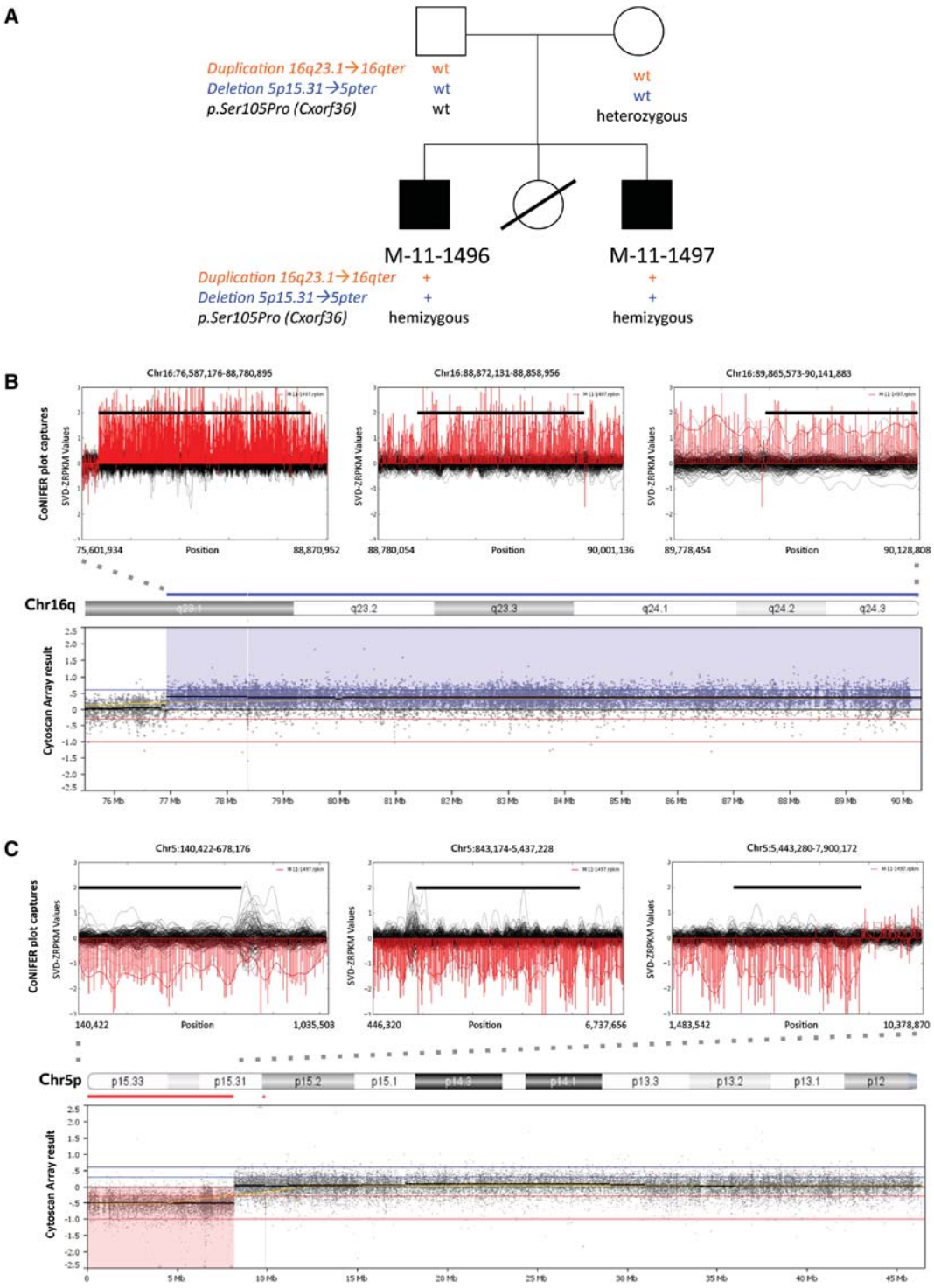


Figure 1. Rare single-nucleotide variants (SNVs) and copy-number variants (CNVs) detected in a Turkish pedigree with a syndrome of unknown etiology. (*A*) Pedigree and genotypes of a nonconsanguineous Turkish family with two affected male siblings. A daughter was presumed to have a different congenital disorder; she had hypotonia and was deceased at 8 mo of age. (*B*) CoNIFER plots, divided into three genomic segments, indicating a breakpoint at position 76,857,179, and a Cytoscan array image (breakpoint at position 76,935,310) are representative of the terminal duplication on Chr 16. (*C*) CoNIFER plots, divided into three genomic segments, indicating a breakpoint at position 7,900,172, and a Cytoscan array image (breakpoint is 8,180,513) are representative of the terminal deletion on Chr 5. Images are from M-11-1497 (see Supplemental Fig. S3 for M-11-1496). Note that the (normal) parental samples are included in the aggregate CoNIFER plots.

For WES, total coverage for each individual ranged from 92% to 94% of bases covered at 10× (Supplemental Table S1). We first filtered variants against 1000 Genomes, and the 6500 exomes in the EVS (Exome Variant Server), retaining only functional variants present at <1% minor allele frequency (MAF). Remaining variants were then filtered against the Atherosclerosis Risk in Communities cohort (ARIC; $n = 2300$ exomes) to obtain 468–476 rare variants per individual (Supplemental Table S2). Next, we performed trio analysis for each sibling separately under de novo, autosomal-recessive, or X-linked hypotheses, followed by cross-referencing across siblings to identify genes that shared the same mutations. Individual M-11-1496 had 11 candidate genes and individual M-11-1497 had nine candidate genes (Supplemental Table S3). Among these putative contributing loci, we found a single shared candidate among the two siblings, a missense mutation on the X Chromosome (c.T313C; p.Ser105Pro) in *CXorf36*. Sanger sequencing confirmed that both siblings were hemizygous for this variant, and that they had each inherited the c.T313C change from their carrier mother; the father was hemizygous reference at this position (Supplemental Fig. S1).

Inactivation in Zebrafish Shows That *CXorf36* Is Not the Likely Driver of Pathogenesis

CXorf36 encodes a 433-amino-acid putative protein with a predicted molecular weight of 48 kDa. The only implication of this locus in human pathology is a tentative association with autism and X-linked intellectual disability (Aziz et al. 2011), whereas the CADD (combined annotation-dependent depletion) score for the discovered allele was 23.8 (and is therefore in the top 1% of likely deleterious mutations [Kircher et al. 2014]). To test this transcript as our sole candidate derived from SNV analysis of the quad, we turned to the developing zebrafish, a system we have used extensively to test functionally candidate genes and alleles (Niederriter et al. 2013) for neurodevelopmental defects and facial dysmorphia (Chassaing et al. 2012; Golzio et al. 2012; Dauber et al. 2013; Brooks et al. 2014). We detected the single ortholog of *CXorf36* in the zebrafish genome (*c9hxorf36*; 48% identity, 89% similarity), and we established that it is expressed during embryonic development by analyzing in-house RNA-seq data from 5 days postfertilization (dpf) heads (Borck et al. 2015). We therefore suppressed *c9hxorf36* using morpholino (MO) antisense oligonucleotides at three different doses (3, 6, and 9 ng). Even at the highest dose, which extinguished ~90% of endogenous mRNA, we did not observe any differences between morphants and control animals (Supplemental Fig. S2). Testing included analysis of the craniofacial skeleton by Alcian blue staining at 5 dpf, followed by measurements of the angle of ceratohyal cartilage and distance between Meckel's cartilage and the ceratohyal cartilage; both yielded measures that were similar between control and MO-injected batches ($n = 40$ embryos/batch; repeated twice; investigator was blind to injection cocktail; Fig. 2A–C). In addition, calculations of the number of dividing cells in the head (to determine the possible relevance to neurodevelopmental defects) revealed no differences in the numbers and distribution of phospho histone-H3 positive cells at 2 dpf (performed in triplicate; Fig. 2D,E). Finally, measurement of head size was also similar between MO-injected embryos versus controls ($n = 35$ embryos/batch; triplicated; Fig. 2F). Similarly, *CXorf36* mRNA overexpression induced no obvious phenotypes. Together, these data suggest that *CXorf36* is unlikely to contribute to the complex clinical presentation of the affected siblings. Consistent with this notion, a recent analysis of exomes in the ExAC (Exome Aggregation Consortium) browser, made available after we had concluded the zebrafish studies, showed that the discovered missense allele, although exceedingly rare in Europeans (five hemizygous males), is present in 91 males of South Asian origin, as well as in two females in homozygosity. In addition, a total of 90 missense and four loss-of-function (LOF) alleles have been reported for this transcript in ExAC, predicting moderate constraint of deleterious alleles for this locus.

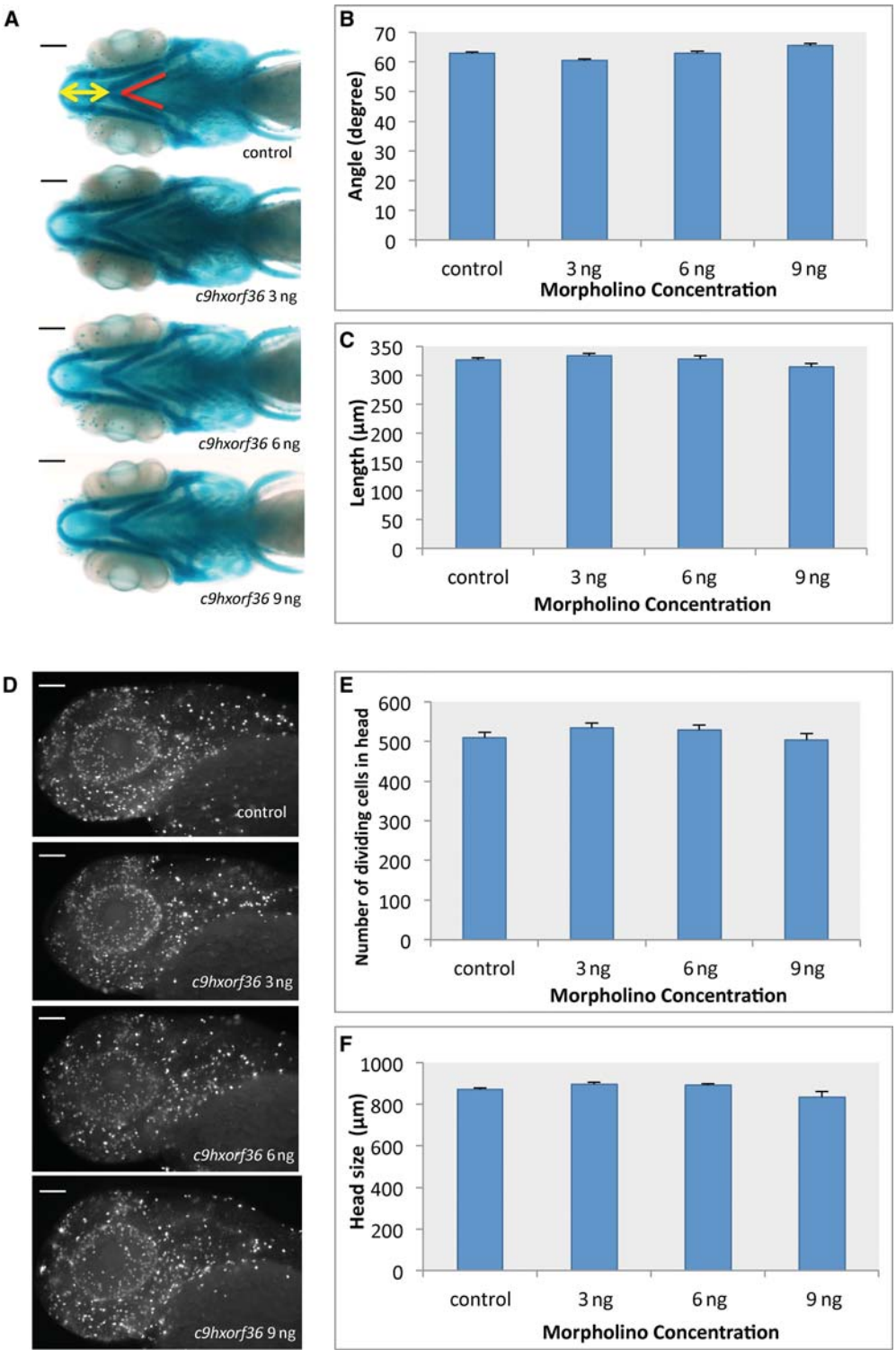


Figure 2. In vivo assessment of *c9hxorf36* suppression in zebrafish. (*A*) Representative ventral images of the craniofacial structure of zebrafish embryos fixed at 5.5 dpf and stained with Alcian blue. (*B*) Quantitative measurement (shown with red lines in *A*) of control and morpholino [MO]-injected batches ($n = 40$ embryos/batch; repeated twice) indicate no significant difference in the angle of ceratohyal cartilage in morphants versus controls. (*Legend continues on following page.*)

CNV Analysis Reveals a 16q Terminal Duplication and a 5p Terminal Deletion

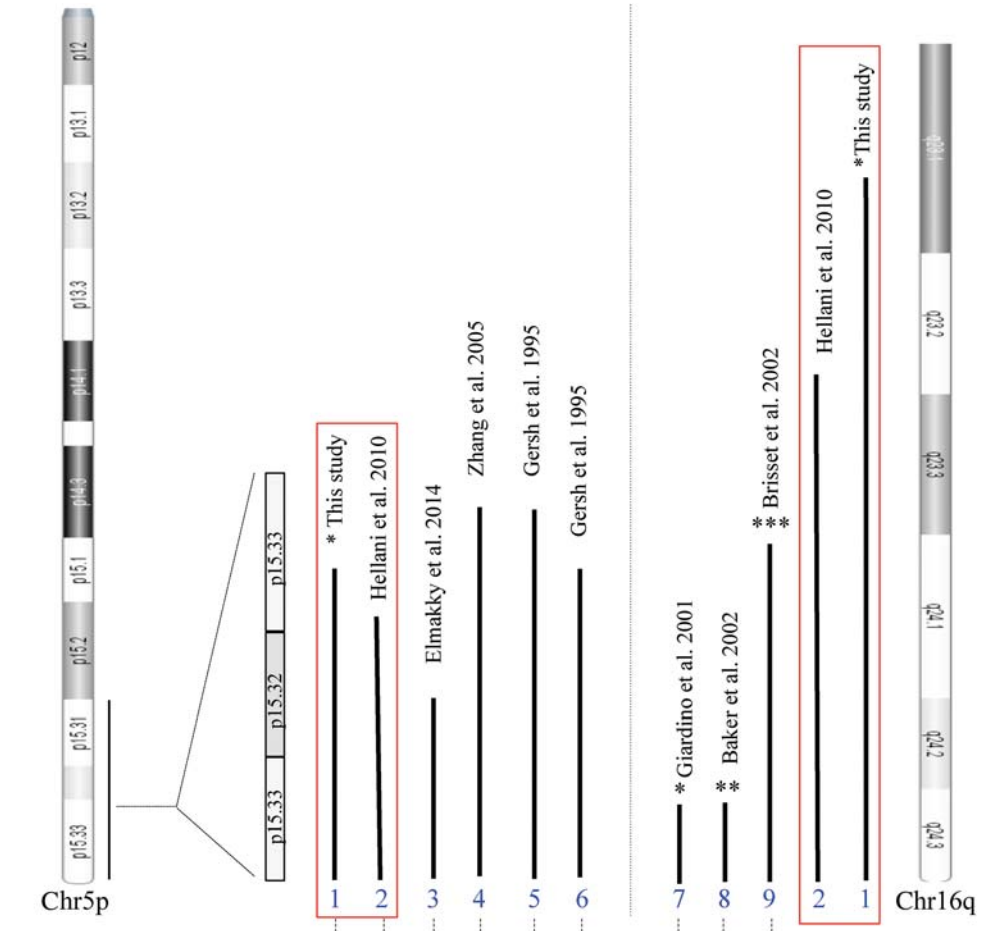
Because our functional data did not support the candidacy of *CXorf36* as a driver, we reexamined our exome data seeking other types of lesions. CNV analysis using CoNIFER (COpy Number Inference From Exome Reads; Krumm et al. 2012) detected two such lesions: a terminal duplication of Chromosome 16q (16q23.1 → 16qter) and a terminal deletion of Chromosome 5p (15p15.31 → 15pter) proximal to the minimal cri du chat region (Figs. 1B, C, 3; Supplemental Fig. S3A,B). Both CNVs were present in both affected siblings and were absent from parental exomes, whereas there are no healthy individuals reported in the Database of Genomic Variants (DGV; http//projects.tcag.ca/variation/) with concomitant losses at 5p and gains at 16q. A second CNV algorithm, XHMM (eXome–Hidden Markov Model; Fromer et al. 2012), reported the same result (Supplemental Table S4). The presence of two CNVs in both children that were undetectable in either parent hinted at a malsegregation of a parental balanced translocation. We therefore performed karyotype analysis of both affected siblings (the parents refused to provide new parental samples), and we confirmed the presence of a t(5;16) translocation.

To study this genomic lesion further, we cross-referenced our exome data with the molecular cytogenetic analysis performed on both patients on a high-resolution single-nucleotide polymorphism (SNP) array. We confirmed both CNVs and we mapped the breakpoints: The terminal deletion at Chromosome 5 starts at p15.31 (8,180,513) between *MTRR* and *LOC7229506* and encompasses a region containing 40 genes, five pseudogenes, and five microRNA genes; the terminal duplication on Chromosome 16 starts at q23.15 (76,935,310) between *MIR4719* and *MON1B* and contains 110 genes, two pseudogenes, and two microRNA genes (Figs. 1B,C and 3). Notably, none of the breakpoints disrupted any transcripts, suggesting that the contribution of these CNVs to the complex clinical presentation would have to be driven by dosage imbalance.

The size (13.5 and 7.7 Mb for 16q and 5p, respectively) and large genic content of each CNV precluded systematic functional in vivo testing. As such, the only tractable interpretive tool was retrospective CNV analysis of affected individuals deposited either in the literature or in public databases such as DGV and the Database of Chromosomal Imbalance and Phenotype in Humans (DECIPHER; https://decipher.sanger.ac.uk). We found several reports of patients with a CNV overlapping partially with either the 5p deletion or the partial 16q trisomy (Fig. 3) and a single report of the co-occurrence of the 5p del and 16q dup in one other family (Hellani et al. 2010).

Despite the expected phenotypic variability associated with the 16q dup and the 5p del, some phenotypic features can be potentially assigned to dosage imbalance at either chromosomal region. For example, the observed terminal deletion of Chromosome 5p overlaps partially with cri du chat syndrome (Zhang et al. 2005), some features of which, such as the cat-like cry speech delay, were recorded in our patient. Similarly, duplications in 16q11 → qter to 16q24 → qter have been reported in patients with features present in our patients, including hypotonia, low-set ears, genital abnormalities, and craniofacial features (Laus et al. 2012). Importantly, a kindred of Saudi Arabian descent with two affected half-brothers

Figure 2. (*Continued*) (*C*) Quantitative measurement (as indicated with yellow arrows in *A*) of the distance between the ceratohyal and Meckel's cartilage resulted in no significant difference between controls and morphant embryos ($n = 40$ embryos/batch; repeated twice). (*D*) Representative lateral images of 2-dpf embryos stained with anti-phospho histone H3 antibody to monitor cell division. (*E*) Quantification of dividing cells in the heads of embryo batches ($n = 15$ embryos/batch); no significant differences were detected. (*F*) Summary of head size measurements as determined by the distance from the most anterior point of Meckel's cartilage to the fin attachment on ventral views (as shown in *A*; $n = 35$ embryos/batch); no significant differences were detected. Scale bars (*A*,*D*), 100 µm.



Number of affected individuals reported in each study (9 studies in total)		2	2	5	1	2	4	1	2	1
Observed phenotypes	Cat-like cry	1/2	NR	5/5	0/1	2/2	3/4	0/1	0/2	0/1
	Speech delay	2/2	2/2	0/5	1/1	0/2	0/4	0/1	0/2	0/1
	Intellectual disability	2/2	2/2	0/5	1/1	0/2	2/4	1/1	0/2	1/1
	Genital hypoplasia	2/2	2/2	NR	NR	NR	NR	0/1	0/2	0/1
	Dysmorphism[1]	2/2	2/2	0/5	0/1	0/2	0/4	0/1	0/2	1/1
	Cardiac defect	0/2	2/2	NR	NR	NR	NR	1/1	0/2	1/1
	Hypotonia	2/2	2/2	2/5	NR	0/2	0/4	1/1	2/2	1/1

Figure 3. Summary of CNVs on Chromosome 5p and Chromosome 16q from this study and previous reports (nine studies in total). *, with 2q37 deletion; **, with 5q35.3 deletion; ***, with 7q22.3 deletion; NR, not recorded. The different reports with CNVs overlapping the 5p and 16q lesions from our study have been assigned a number (1–9, shown in blue). Each row of observed phenotypes indicates the number of individuals with that feature per number evaluated in each report. Phenotypes in bold are both concordant among the individuals with concomitant 5p and 16q CNVs from reports 1 (this study) and 2 (Hellani et al. 2010). [1]Facial dysmorphism includes elongated face with prominent forehead, broad nasal bridge, low ear sets, arched palate, and thin upper lip.

was shown to harbor cytogenetic changes and clinical features that overlapped with those in our family: a 1-Mb 5p terminal deletion and a 6-Mb 16q terminal duplication. These two affected siblings bore striking similarities with our cases with respect to their neurodevelopmental and genital defects. Similar to our family, the most parsimonious mechanism is the missegregation of a balanced translocation (Table 1; Fig. 3; Hellani et al. 2010).

DISCUSSION

Here, we have reported our genetic analyses of a nonconsanguineous Turkish pedigree with a constellation of structural anomalies affecting neurodevelopment, craniofacial and skeletal formation, and the genitourinary tract that do not conform readily to a recognized syndrome. Using a quad-based WES strategy combined with high-resolution SNP arrays, we found three candidate pathogenic drivers: a missense mutation on the X Chromosome and two CNVs arising from a likely balanced translocation in one of the parents. Functional testing of *CXorf36* revealed that the mutation is unlikely to be contributory, at least for the neurodevelopmental and craniofacial defects tested.

Retrospective review of cases with 5p deletions and 16q duplications that overlap the regions found in our patients suggest that each chromosomal aberration likely contributes to the complexity of the phenotype, with features such as the cat-like cry in one of our patients arising from the 5p deletion and hypotonia being shared as a consistent feature of the 16q duplication (Fig. 3). At the same time, we note that neither 5p (del) nor 16q (dup) have been associated with genital hypoplasia. Although this might be a product of variable expressivity at either CNV, this is a feature shared between our patients and those reported to have both 5p and 16q dosage imbalances (Hellani et al. 2010). The 16q duplication in both these studies is larger than what has been reported previously. Thus, one possibility is that the increased expression of transcripts in that interval drives this pathology. Alternatively, we speculate that genetic interactions between the two CNVs might be a driver of the structural malformations of the male gonads. Finally, given that our studies were focused primarily on the exome, we cannot exclude the possibility that additional pathogenic changes (SNV or CNV; regulatory or coding) elsewhere in the genome contribute to pathology, an increasingly reported phenomenon coincident with the expanded use of whole genome sequencing (Lelieveld et al. 2015; Lupianez et al. 2015). The identification and careful phenotyping for each of these pathologies of additional patients harboring either or both CNVs as well as comprehensive whole-genome analysis of the individuals in this study will be necessary to resolve these possibilities.

We do not yet know which genes within these two CNVs contribute to each of the observed phenotypes; ultimately, the systematic dissection of CNVs will be required to identify phenotype drivers and potential epistatic effects resulting from genes affected by CNV(s) detected in humans (Golzio et al. 2012; Dauber et al. 2013). We anticipate that the unbiased dissection of large CNVs similar in size to those identified in this study will become possible with the increased throughput of genome-editing technologies (Hsu et al. 2014) in model organisms.

METHODS

Ascertainment, WES, CoNIFER and XHMM Analysis, and Sanger Sequencing

The parents and two affected siblings were sequenced using WES on an Illumina HiSeq platform using 100-bp paired-end reads at a median depth of >80% at 20× as described previously (Bainbridge et al. 2011). Briefly, DNA samples were fragmented and ligated with

Illumina multiplex paired-end adapters. Adapter-ligated fragmented DNA was amplified using primers with sequencing barcodes. The precapture library was enriched by hybridizing to biotin-labeled VCRome 2.1 (Bainbridge et al. 2011). Output data were mapped by Burrows–Wheeler aligner (BWA) (Li and Durbin 2009). The variant calls were performed with Atlas-SNP and Atlas-indel as described previously (Challis et al. 2012). Data were filtered to retain functional variants (predicted to alter mRNA splicing or protein amino acid sequence) with an MAF of <1% in the public 1000 Genomes SNP database, the National Heart, Lung, and Blood Institute (NHLBI) Exome Sequencing Project, the Exome Variant Server (EVS; n = 6500 exomes), and the ARIC cohort (n = 2300 exomes). Finally, candidate variants were confirmed visually using the Integrative Genomics Viewer (IGV), followed by Sanger sequencing. Retained variants were placed into two categories: (1) primary candidates—variants that segregate perfectly with disease status, with coverage ≥5× and allele balance ≥20% in the proband; and (2) secondary candidates—inheritance is consistent with disease status (genotype of one parent may be missing; coverage may be <5× in one of the parents; reduced quality [e.g., more than two alleles at one site, one of which has <20% of reads]; coverage is ≥5× in the proband; see Supplemental Tables S1–S3 for details).

CNVs were identified using two programs: CoNIFER (Krumm et al. 2012) and XHMM (Fromer et al. 2012). Both programs were run with default parameters. To ensure highest sensitivity, we did not perform any filtering of the discovered CNVs. In addition to the four samples from the family, we included WES data from unrelated 200 in-house samples to enable better normalization of the depth of coverage. The discovered CNVs in each proband were annotated using a custom Perl script indicating the amount of overlap with known segmental duplications and known CNVs from the Database of Genomic Variants (DGV) as reported in the UCSC Genome Browser (http://genome.ucsc.edu/cgi/hgGateway).

High-Resolution Array

Array analysis was carried out following the manufacturer's protocols on the high-resolution CytoScan HD array platform (Affymetrix), which contains more than 2.6 million markers, including 750,000 SNPs and 1.9 million nonpolymorphic probes. Data were analyzed with the Nexus software version 7.5 (BioDiscovery). We utilized the two major quality control metrics for Affymetrix array: (1) the median absolute pairwise difference (MAPD) score, which applies to copy-number probes, and (2) the SNP-quality control (QC), which applies to SNP probes. We mandated that the values for these parameters be ≤0.25 for MAPD and ≥0.15 for SNP-QC. We set cutoffs for our detection criteria for CNVs at 20 probes for both gains and losses, and the breakpoint positions of aberrant regions were converted to UCSC hg19 (UCSC Genome Browser, release February 2009).

Zebrafish Embryo Injection and Phenotyping

A splice-blocking MO targeting the exon 2–intron 2 boundary of *c9hxorf36* was designed and obtained from Gene Tools (5′-AAGCCACAAATTCAGACCTTCACCA-3′). The efficiency of the MO was confirmed by RT-PCR (reverse transcription–polymerase chain reaction). RNA was extracted from MO-injected embryos and controls at 3 dpf using TRIzol (Invitrogen). First-strand cDNA was synthesized using SuperScript III reverse transcriptase and random hexamers (Invitrogen), according to the manufacturer's instructions. We carried out PCR amplification using whole-embryo cDNA as template, and PCR products were subjected to direct DNA sequencing (Supplemental Fig. S2). One nanoliter of diluted MO (3, 6, and 9 ng) was injected into zebrafish embryos at the one- to four-cell stage. Injected embryos were fixed in 4% PFA (paraformaldehyde) overnight at 5.5 dpf (Alcian blue staining) and at 2 dpf for phospho-histone H3 staining. Alcian blue staining was performed to demarcate cartilage structures as described previously (Chassaing et al. 2012). Phospho-histone H3

immunostaining was performed as described previously (Golzio et al. 2012) using anti-histone H3 (ser10)-R, (sc-8656-R, Santa Cruz; 1:500 dilution). Embryos were imaged on a Nikon AZ100 microscope facilitated by NIS Elements software.

ADDITIONAL INFORMATION

Ethics Statement

All samples were collected following written informed consent under the requirements and with the approval of Hacettepe University Academic ethics committee and the Duke University Institutional Review Board.

Database Deposition and Access

The t(5;16) translocation reported in this study has been deposited in ClinVar (http://www.ncbi.nlm.nih.gov/clinvar/) under accession number SCV000257401. Consent could not be obtained permitting the public deposition of whole-exome sequencing data.

Acknowledgments

We are grateful to this family for their interest in and support of our work.

Author Contributions

A.O., E.E.D., and N.K. designed the study; K.R.Ö., D.Y., E.S., and A.D. collected patient sample and clinical information; A.O., A.S., D.M., S.D.-P., and R.A.G. performed whole-exome analysis; M.M.W. and E.A.S. performed SNP array analysis; A.O. and E.E.D. performed zebrafish analyses; A.O., E.E.D., A.D., and N.K. wrote the manuscript.

Competing Interest Statement

The authors have declared no competing interest.

Received August 30, 2015; accepted in revised form November 23, 2015.

Funding

This work was funded by the Simons Foundation Autism Research Initiative (SFARI) 239983, and U.S. National Institutes of Health (NIH) grants P50DK096415 (N.K.) and U54HG003273 (R.A.G.). N.K. is a Distinguished George W. Brumley Professor.

REFERENCES

Aziz A, Harrop SP, Bishop NE. 2011b. *DIA1R* is an X-linked gene related to *Deleted in Autism-1*. *PLoS One* **6:** e14534.

Bainbridge MN, Wang M, Wu Y, Newsham I, Muzny DM, Jefferies JL, Albert TJ, Burgess DL, Gibbs RA. 2011. Targeted enrichment beyond the consensus coding DNA sequence exome reveals exons with higher variant densities. *Genome Biol* **12:** R68.

Baker E, Hinton L, Callen DF, Altree M, Dobbie A, Eyre HJ, Sutherland GR, Thompson E, Thompson P, Woollatt E, et al. 2002. Study of 250 children with idiopathic mental retardation reveals nine cryptic and diverse subtelomeric chromosome anomalies. *Am J Med Genet* **107:** 285–293.

Beaulieu CL, Majewski J, Schwartzentruber J, Samuels ME, Fernandez BA, Bernier FP, Brudno M, Knoppers B, Marcadier J, Dyment D, et al. 2014. FORGE Canada Consortium: outcomes of a 2-year national rare-disease gene-discovery project. *Am J Hum Genet* **94:** 809–817.

Borck G, Hog F, Dentici ML, Tan PL, Sowada N, Medeira A, Gueneau L, Thiele H, Kousi M, Lepri F, et al. 2015. BRF1 mutations alter RNA polymerase III–dependent transcription and cause neurodevelopmental anomalies. *Genome Res* **25:** 155–166.

Brisset S, Joly G, Ozilou C, Lapierre JM, Gosset P, LeLorc'h M, Raoul O, Turleau C, Vekemans M, Romana SP. 2002. Molecular characterization of partial trisomy 16q24.1-qter: clinical report and review of the literature. *Am J Med Genet* **113:** 339–345.

Brooks SS, Wall AL, Golzio C, Reid DW, Kondyles A, Willer JR, Botti C, Nicchitta CV, Katsanis N, Davis EE. 2014. A novel ribosomopathy caused by dysfunction of RPL10 disrupts neurodevelopment and causes X-linked microcephaly in humans. *Genetics* **198:** 723–733.

Challis D, Yu J, Evani US, Jackson AR, Paithankar S, Coarfa C, Milosavljevic A, Gibbs RA, Yu F. 2012. An integrative variant analysis suite for whole exome next-generation sequencing data. *BMC Bioinformatics* **13:** 8.

Chassaing N, Sorrentino S, Davis EE, Martin-Coignard D, Iacovelli A, Paznekas W, Webb BD, Faye-Petersen O, Encha-Razavi F, Lequeux L, et al. 2012. *OTX2* mutations contribute to the otocephaly-dysgnathia complex. *J Med Genet* **49:** 373–379.

Dauber A, Golzio C, Guenot C, Jodelka FM, Kibaek M, Kjaergaard S, Leheup B, Martinet D, Nowaczyk MJ, Rosenfeld JA, et al. 2013. SCRIB and PUF60 are primary drivers of the multisystemic phenotypes of the 8q24.3 copy-number variant. *Am J Hum Genet* **93:** 798–811.

Davis EE, Katsanis N. 2012. The ciliopathies: a transitional model into systems biology of human genetic disease. *Curr Opin Genet Dev* **22:** 290–303.

Elmakky A, Carli D, Lugli L, Torelli P, Guidi B, Falcinelli C, Fini S, Ferrari F, Percesepe A. 2014. A three-generation family with terminal microdeletion involving 5p15.33-32 due to a whole-arm 5;15 chromosomal translocation with a steady phenotype of atypical cri du chat syndrome. *Eur J Med Genet* **57:** 145–150.

Fromer M, Moran JL, Chambert K, Banks E, Bergen SE, Ruderfer DM, Handsaker RE, McCarroll SA, O'Donovan MC, Owen MJ, et al. 2012. Discovery and statistical genotyping of copy-number variation from whole-exome sequencing depth. *Am J Hum Genet* **91:** 597–607.

Gersh M, Goodart SA, Pasztor LM, Harris DJ, Weiss L, Overhauser J. 1995. Evidence for a distinct region causing a cat-like cry in patients with 5p deletions. *Am J Hum Genet* **56:** 1404–1410.

Giardino D, Finelli P, Gottardi G, Clerici D, Mosca F, Briscioli V, Larizza L. 2001. Cryptic subtelomeric translocation t(2;16)(q37;q24) segregating in a family with unexplained stillbriths and a dysmorphic, slightly retarded child. *Eur J Hum Genet* **9:** 881–886.

Girirajan S, Rosenfeld JA, Coe BP, Parikh S, Friedman N, Goldstein A, Filipink RA, McConnell JS, Angle B, Meschino WS, et al. 2012. Phenotypic heterogeneity of genomic disorders and rare copy-number variants. *N Engl J Med* **367:** 1321–1331.

Golzio C, Willer J, Talkowski ME, Oh EC, Taniguchi Y, Jacquemont S, Reymond A, Sun M, Sawa A, Gusella JF, et al. 2012. KCTD13 is a major driver of mirrored neuroanatomical phenotypes of the 16p11.2 copy number variant. *Nature* **485:** 363–367.

Hellani A, Mohamed S, Al-Akoum S, Bosley TM, Abu-Amero KK. 2010. A t(5;16)(p15.32;q23.3) generating 16q23.3 → qter duplication and 5p15.32 → pter deletion in two siblings with mental retardation, dysmorphic features, and speech delay. *Am J Med Genet A* **152A:** 1555–1560.

Hsu PD, Lander ES, Zhang F. 2014. Development and applications of CRISPR-Cas9 for genome engineering. *Cell* **157:** 1262–1278.

Kircher M, Witten DM, Jain P, O'Roak BJ, Cooper GM, Shendure J. 2014. A general framework for estimating the relative pathogenicity of human genetic variants. *Nat Genet* **46:** 310–315.

Krumm N, Sudmant PH, Ko A, O'Roak BJ, Malig M, Coe BP; NHLBI Exome Sequencing Project, Quinlan AR, Nickerson DA, Eichler EE. 2012. Copy number variation detection and genotyping from exome sequence data. *Genome Res* **22:** 1525–1532.

Laus AC, Baratela WA, Laureano LA, Santos SA, Huber J, Ramos ES, Rebelo CC, Squire JA, Martelli L. 2012. Karyotype/phenotype correlation in partial trisomies of the long arm of chromosome 16: case report and review of literature. *Am J Med Genet A* **158A:** 821–827.

Lelieveld SH, Spielmann M, Mundlos S, Veltman JA, Gilissen C. 2015. Comparison of exome and genome sequencing technologies for the complete capture of protein-coding regions. *Hum Mutat* **36:** 815–822.

Lemmers RJ, Tawil R, Petek LM, Balog J, Block GJ, Santen GW, Amell AM, van der Vliet PJ, Almomani R, Straasheijm KR, et al. 2012. Digenic inheritance of an *SMCHD1* mutation and an FSHD-permissive D4Z4 allele causes facioscapulohumeral muscular dystrophy type 2. *Nat Genet* **44:** 1370–1374.

Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25:** 1754–1760.

Lindstrand A, Davis EE, Carvalho CM, Pehlivan D, Willer JR, Tsai IC, Ramanathan S, Zuppan C, Sabo A, Muzny D, et al. 2014. Recurrent CNVs and SNVs at the NPHP1 locus contribute pathogenic alleles to Bardet–Biedl syndrome. *Am J Hum Genet* **94:** 745–754.

Lupianez DG, Kraft K, Heinrich V, Krawitz P, Brancati F, Klopocki E, Horn D, Kayserili H, Opitz JM, Laxova R, et al. 2015. Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. *Cell* **161:** 1012–1025.

Ng SB, Buckingham KJ, Lee C, Bigham AW, Tabor HK, Dent KM, Huff CD, Shannon PT, Jabs EW, Nickerson DA, et al. 2010. Exome sequencing identifies the cause of a Mendelian disorder. *Nat Genet* **42:** 30–35.

Niederriter AR, Davis EE, Golzio C, Oh EC, Tsai IC, Katsanis N. 2013. In vivo modeling of the morbid human genome using *Danio rerio*. *J Vis Exp* e50338.

Roach JC, Glusman G, Smit AF, Huff CD, Hubley R, Shannon PT, Rowen L, Pant KP, Goodman N, Bamshad M, et al. 2010. Analysis of genetic inheritance in a family quartet by whole-genome sequencing. *Science* **328:** 636–639.

Schaaf CP, Wiszniewska J, Beaudet AL. 2011. Copy number and SNP arrays in clinical diagnostics. *Annu Rev Genomics Hum Genet* **12:** 25–51.

Snijders AM, Nowak N, Segraves R, Blackwood S, Brown N, Conroy J, Hamilton G, Hindle AK, Huey B, Kimura K, et al. 2001. Assembly of microarrays for genome-wide measurement of DNA copy number. *Nat Genet* **29:** 263–264.

Solinas-Toldo S, Lampel S, Stilgenbauer S, Nickolenko J, Benner A, Dohner H, Cremer T, Lichter P. 1997. Matrix-based comparative genomic hybridization: biochips to screen for genomic imbalances. *Genes Chromosomes Cancer* **20:** 399–407.

Yang Y, Muzny DM, Reid JG, Bainbridge MN, Willis A, Ward PA, Braxton A, Beuten J, Xia F, Niu Z, et al. 2013. Clinical whole-exome sequencing for the diagnosis of Mendelian disorders. *N Engl J Med* **369:** 1502–1511.

Zhang X, Snijders A, Segraves R, Zhang X, Niebuhr A, Albertson D, Yang H, Gray J, Niebuhr E, Bolund L, et al. 2005. High-resolution mapping of genotype–phenotype relationships in cri du chat syndrome using array comparative genomic hybridization. *Am J Hum Genet* **76:** 312–326.