

ARTICLE

Clinical and genomic characterization of 8p cytogenomic disorders

Volkan Okur^{1,2}, Laura Hamm¹, Haluk Kavus¹, Caroline Mebane¹, Scott Robinson¹, Brynn Levy³ and Wendy K. Chung^{1,4}✉

PURPOSE: To provide a detailed clinical and cytogenomic summary of individuals with chromosome 8p rearrangements of invdupdel(8p), del(8p), and dup(8p).

METHODS: We enrolled 97 individuals with invdupdel(8p), del(8p), and dup(8p). Clinical and molecular data were collected to delineate and compare the clinical findings and rearrangement breakpoints. We included additional 5 individuals with dup(8p) from the literature for a total of 102 individuals.

RESULTS: Eighty-one individuals had recurrent rearrangements of invdupdel(8p) ($n = 49$), del(8p)_distal ($n = 4$), del(8p)_proximal ($n = 9$), del(8p)_proximal&distal ($n = 12$), and dup(8p)_proximal ($n = 7$). Twenty-one individuals had nonrecurrent rearrangements. While all individuals had neurodevelopmental features, the frequency and severity of clinical findings were higher in individuals with invdupdel(8p), and with larger duplications. All individuals with *GATA4* deletion had structural congenital heart defects; however, the presence of structural heart defects in some individuals with normal *GATA4* copy number suggests there are other potentially contributing gene(s) on 8p.

CONCLUSION: Our study may inform families and health-care providers about the associated clinical findings and severity in individuals with chromosome 8p rearrangements, and guide researchers in investigating the underlying molecular and biological mechanisms by providing detailed clinical and cytogenomic information about individuals with distinct 8p rearrangements.

Genetics in Medicine (2021) 23:2342–2351; <https://doi.org/10.1038/s41436-021-01270-2>

INTRODUCTION

Genomic disorders, i.e., deletions and duplications, account for 10–15% of congenital genetic diseases including multiple congenital anomalies, neurodevelopmental, and neurobehavioral disorders [1, 2]. Recurrent rearrangements of chromosome 8p were first reported ~45 years ago by conventional cytogenetics [3]. Advancements in cytogenomics have enabled refinement of chromosome 8p rearrangements.

The short arm of chromosome 8 (8p) contains low copy repeat (LCR) regions such as olfactory receptor and defensin gene clusters, and parental inversion can mediate nonallelic homologous recombination (NAHR) leading to an inverted duplication of the interstitial region and deletion of the telomeric region of 8p, known as inverted duplication/deletion 8p (invdupdel[8p]) with variability in duplication size, deletion of telomeric and interstitial regions of 8p (del[8p]), and duplication of interstitial regions of 8p (dup[8p]) [4].

Rearrangements of chromosome 8p can span 80–90% of the entire short arm (~45 Mb), particularly in invdupdel(8p), and the phenotype may be due to one or more dosage sensitive genes. Our goal was to delineate the breakpoints of recurrent 8p rearrangements and compare clinical phenotypes across individuals to better support the care of individuals with 8p disorders and identify individuals with small copy-number variants (CNVs) who might provide insight into critical genes on 8p contributing to specific phenotypes.

MATERIALS AND METHODS

Patient ascertainment and data collection

We recruited individuals with cytogenetically (chromosome analysis, fluorescence in situ hybridization [FISH], or chromosome microarray [CMA]) documented invdupdel(8p), del(8p), or dup(8p) without other pathogenic CNVs. We collected and reviewed the genetic reports and clinical records when available. Medical history interviews and Vineland Adapted Behavior Scales, Third Edition (Vineland-3) were conducted with primary caregivers. Facial photos were collected with additional consent for publication of photos. When genomic coordinates of rearrangements were not available, CMA studies were performed in a clinical diagnostic laboratory using genomic DNA extracted from saliva or blood to delineate the breakpoints for individuals with only prior chromosome analysis and/or FISH.

To the best of our knowledge, only two individuals (siblings) with dup(8p) from our cohort have been previously published [5]. We did not include literature reports of previously reported individuals with invdupdel(8p) and del(8p) to avoid double counting of individuals. We added five individuals with recurrent dup(8p) from the published literature to achieve representative number of individuals [6, 7].

Data analysis

Genomic findings. Since the CMA studies were performed as routine clinical care using different platforms, all breakpoint coordinates were converted to GRCh37/hg19 using the University of California–Santa Cruz (UCSC) LiftOver Tool (<http://genome.ucsc.edu>). For gene dosage metrics, we used probability of loss of function intolerance (pLI) and loss-of-function observed/expected upper bound fraction (LOEUF) metric provided by gnomAD [8] and updated haploinsufficiency percentage

¹Division of Molecular Genetics, Department of Pediatrics, Columbia University Medical Center, New York, NY, USA. ²Department of Human and Molecular Genetics, Baylor College of Medicine, Houston, TX, USA. ³Department of Pathology and Cell Biology, Columbia University Medical Center, New York, NY, USA. ⁴Department of Medicine, Columbia University Medical Center, New York, NY, USA. ✉email: wkc15@cumc.columbia.edu

Table 1. Demographic and clinical characteristics of individuals with recurrent 8p rearrangements.

Category	Invdupdel(8p) (n = 49)	del(8p)_distal (n = 4)	del(8p)_proximal&distal (n = 9)	del(8p)_proximal (n = 12)
Shared intervals				
Deletion coordinates	1–6,697,895	1–6,746,449	928,898–10,767,748	9,544,460–11,498,149
Duplication coordinates	12,635,000–23,675,945	NA	NA	NA
Age distribution (years)				
Mean age	7.1	18.6	16.2	7.3
Median age	5.5	16.1	11.5	4.8
Youngest–oldest participant	0.9–27.2	8.3–33.9	0.8–32.1	1.7–17.2
Prenatal findings				
	47%	25%	44%	75%
Single umbilical artery	16%	-	-	17%
SGA	14%	-	33%	25%
Cardiac finding	10%	-	-	25%
Oligo-/polyhydramnios, renal issues	8%	25%	-	-
Neonatal issues				
	92%	50%	78%	100%
Floppy infant	71%	25%	44%	25%
Feeding difficulty	63%	-	22%	42%
Poor suck	57%	-	33%	42%
Overly sleepy	43%	50%	11%	33%
Irritability	27%	-	22%	58%
Prematurity	22%	-	11%	17%
Respiratory distress	18%	-	22%	17%
Hypoglycemia	18%	25%	22%	8%
Anthropometric measurements				
Birth weight Z-scores				
Mean	-0.2	0.08	-1.0	-1.45
Median	-0.13	0.3	-0.8	-1.38
25–75%	-0.85–0.48	-0.85–1.20	-1.58 – -0.28	-2.56 – -0.48
Birth length Z-scores				
Mean	0.25	0.32	-0.16	-0.85
Median	0.48	-0.27	-0.06	-1.15
25–75%	-0.8–1.4	-1.13–1.19	-1–0.85	-2.05–0.06
Birth OFC Z-scores				
Mean	0.1	-0.27	-1.49	-1.44
Median	0.1	-0.57	-1.27	-1.43
25–75%	-0.7–0.9	-1.35–0.51	-2.63 – -0.51	-2.17 – -0.76
Developmental milestones (months)				
Age at sitting				
Mean	17.3	7	7.4	8.7
Median	12	7	8	11
25–75%	9–24	6–8	6–8	6–11
Age at first words				
Mean	29	16	21	21
Median	24	14	18	22
25–75%	20–45	14–16	15–28	12–26
Age at walking				
Mean	41	18	18	19
Median	37	18	18	18
25–75%	24–48	15–21	15–22	15–20
Growth and endocrine issues				
Short stature (ISD range)	12% ([-2.44] – [-3.69])	25% ([-2.0])	11% ([-2.51])	17% ([-2.5] – [-2.9])
FTT/poor weight gain	37%	-	11%	17%
Obesity/overweight	-	50%	67%	30%

Table 1 continued				
Category	Invdupdel(8p) (n = 49)	del(8p)_distal (n = 4)	del(8p)_proximal&distal (n = 9)	del(8p)_proximal (n = 12)
Neurologic findings	90%	75%	100%	92%
Hypotonia	88%	50%	56%	42%
Hypertonia	39%	25%	56%	25%
Difficulty walking/clumsiness	29%	-	67%	33%
Macrocephaly	10%	-	-	-
Microcephaly	4%	-	22%	58%
Seizures	55%	25%	44%	42%
Absence	37%	-	11%	25%
Febrile	18%	-	-	8%
Tonic-clonic	12%	-	11%	8%
Atonic	8%	-	-	8%
Myoclonic	4%	25%	-	-
Abnormal EEG	53% (n = 19/36)	33% (n = 1/3)	67% (n = 2/3)	20% (n = 1/5)
Brain imaging (MRI/CT) abnormality	84% (n = 38/45)	50% (n = 1/2)	40% (n = 2/5)	40% (n = 2/5)
Corpus callosum hypoplasia/aplasia	67%	50%	-	-
Hydrocephalus/ventriculomegaly	27%	-	-	20%
Cerebral/cerebellar atrophy	16%	-	-	-
Dandy-Walker	9%	-	-	-
Intracranial cyst	9%	-	-	-
Other	22%	-	40%	20%
Neurobehavioral issues	43%	100%	88% (n = 7/8)	75%
Stereotype	22%	25%	13%	25%
Aggressivity/tantrums/impulsivity	14%	-	50%	33%
Sensory issues	8%	25%	-	25%
Autism	6%	50%	25%	17%
ADD/ADHD/hyperactivity	4%	50%	38%	58%
Echolalia	2%	-	11%	17%
Anxiety	2%	25%	13%	8%
Cardiovascular findings	65% (n = 26/40)	25% (n = 1/4)	86% (n = 6/7)	100% (n = 11/11)
VSD	30%	-	71%	36%
ASD	15%	-	57%	100%
PFO	18%	-	-	-
PDA	13%	-	29%	-
PS	3%	25%	71%	64%
Arrhythmia	-	-	33% (n = 3/9)	8% (n = 1/12)
Gastrointestinal issues	88%	75%	100%	83%
Constipation	73%	25%	67%	58%
GERD	55%	50%	56%	33%
Feeding difficulty	22%	-	-	17%
Diarrhea	6%	-	11%	8%
Musculoskeletal issues	59%	75%	22%	58%
Scoliosis	22%	25%	-	25%
Vertebral abnormalities	8%	25%	-	-
Hypermobility joints	8%	-	11%	8%
Coxa valga/hip anomalies	8%	-	-	17%
Rib anomalies	6%	-	-	-
Pes planus	4%	25%	11%	17%
Equinovarus/equinovagis	8%	-	22%	8%
Inguinal hernia	8%	50%	-	-
Visual issues	55%	75%	56%	50%
Refractive errors	29%	75%	33%	50%
Strabismus	18%	-	33%	17%
Cortical visual impairment	12%	-	11%	-
Optic nerve (atrophy/enlarged/cupped)	8%	-	-	-

Table 1 continued

Category	Invdupdel(8p) (n = 49)	del(8p)_distal (n = 4)	del(8p)_proximal&distal (n = 9)	del(8p)_proximal (n = 12)
Genitourinary issues	33%	50%	22%	17%
Cryptorchidism	35% (n = 7/20)	25% (n = 1/4)	-	-
Hypospadias	10% (n = 2/20)	-	25% (n = 1/4)	20% (n = 1/5)
Hydronephrosis, kidney issues	12%	50%	-	-
Frequent infections	43%	25%	22%	25%
Otitis media	33%	25%	11%	25%
Upper respiratory	16%	-	11%	17%
Urinary tract	14%	-	11%	17%
Sleep problems	45%	25%	22%	58%
Dental problems	53%	50%	56%	33%
Skin problems	45%	25%	22%	25%

ADD attention deficit disorder, ADHD attention deficit–hyperactivity disorder, ASD atrial septal defect, CT computed tomography, EEG electroencephalogram, FTT failure to thrive, GERD gastroesophageal reflux disease, MRI magnetic resonance imaging, NA not applicable, PDA patent ductus arteriosus, PFO patent foramen ovale, PS pulmonary stenosis, SD standard deviation, SGA small for gestational age, VSD ventricular septal defect.

scores (HI%) provided by DECIPHER (<https://decipher.sanger.ac.uk/>) [9]. Thresholds of $pLI \geq 0.8$ and $HI\% \leq 10\%$ were used for evaluating genic constraint against loss-of-function variation and haploinsufficiency, respectively.

The breakpoints of the deletions and duplications are given as proximal or distal according to their positions relative to the centromere; i.e., proximal for the breakpoint closer to the centromere.

Clinical findings. For anthropometric values, we used PediTools (<https://peditools.org/>) to calculate adjusted percentile and Z-scores [10]. We used Fenton [11] and World Health Organization (WHO) growth charts for preterm individuals. We also selected WHO growth charts for term babies up to 2 years of age. Between 2 and 20 years of age, we used Centers for Disease Control and Prevention (CDC) growth charts for all individuals. We report the mean, median, and 25–75th percentile of Z-score values for birth weight, length, and occipitofrontal circumference (OFC) of each rearrangement group. For individuals over 20 years of age, we also used the CDC 2–20 years of age growth charts to compare an individual's height with a 20-year-old sex-matched control. CDC body mass index (BMI) calculation tool (<https://www.cdc.gov/healthyweight/bmi/calculator.html>) was used to determine whether an individual was overweight (BMI for age in children = 85–95%; in adults = 25–29.9) or obese (BMI for age in children >95%; in adults ≥ 30).

We report mean, median, and 25–75th percentile values of ages in months for achieving developmental milestones of sitting, walking, first words, and first sentences. We report the number and mean age of individuals who have not achieved each milestone. Individuals who had not reached the youngest age when a given milestone is expected to be achieved, i.e., 12 months of age for first words and first steps, were not included in the calculations.

For categorical values, we report fraction of individuals with a given clinical finding. We also provide characteristics of certain clinical findings such as seizures, structural brain abnormalities, and cardiac defects.

RESULTS

Demographics

We enrolled 97 individuals with documented chromosome 8p rearrangements. The majority of the participants were from the United States ($n = 53$), the United Kingdom ($n = 14$), and other European countries ($n = 20$). We report on 52 individuals with invdupdel(8p), 37 with del(8p), and 8 with dup(8p) (Table S1). Demographic and clinical characteristics of each rearrangement group are summarized in Table 1, and detailed data are provided in Table S1.

Molecular findings

Deletion/duplication patterns. We used the UCSC Genome Browser to visualize and compare the breakpoints of each individual within each rearrangement group (Fig. 1). We attribute the small variations (~0.1–0.3 Mb) between similar breakpoint patterns to variations in microarray versions and platforms, and difficulty accurately assessing sizes of CNVs associated with LCRs. Larger (~0.5–1.5 Mb) variations of breakpoints for recurrent deletions and/or duplications can be attributable to large LCR regions spanning all breakpoints.

Eighty-one individuals had recurrent rearrangements of invdupdel(8p) ($n = 49$), del(8p)_distal ($n = 4$), del(8p)_proximal&distal ($n = 9$), del(8p)_proximal ($n = 12$), and dup(8p) ($n = 7$); the latter including the five individuals from the literature. The detailed description of breakpoint coordinates and visual representation of each recurrent and nonrecurrent ($n = 21$) rearrangement is given in Supplementary Materials (Table S1 and Figs. S1–3). Briefly, the shared deleted interval in individuals with invdupdel(8p) was the most distal 6.7 Mb and the shared duplicated interval was approximately 11.1 Mb. Individuals with del(8p) were grouped into three subcategories based on the location of their deletions: (1) del(8p)_distal, (2) del(8p)_proximal, (3) del(8p)_proximal&distal (Fig. 1). The shared deleted intervals were 6.7 Mb, 3.8 Mb, and 10.7 Mb for del(8p)_distal, del(8p)_proximal, and del(8p)_proximal&distal. The shared duplicated interval was 3.7 Mb in individuals with dup(8p)_proximal.

Gene content and candidate gene prioritization. Chromosome 8p contains ~250 protein-coding genes, and almost every gene is included in at least one of the recurrent rearrangements. The detailed information such as chromosome 8p coordinates, gene constraint metrics, OMIM disease entries, gene functions, and animal models for each gene can be found in Table S2. Briefly, we prioritized the candidate genes *DLGAP2* and *CSMD1* for del(8p)_distal and the deleted segment of invdupdel(8p), *GATA4* and *XKR6* for del(8p)_proximal and dup(8p)_proximal, and *RHOBTB2* and *CHRNA2* for the duplicated segment of invdupdel(8p) (Supplemental Data).

Clinical findings of recurrent cohort. Clinical findings of the recurrent cohort are summarized and compared among subgroups in Table 1, and the most prominent findings by organ systems are outlined in the following sections. Detailed data on clinical findings of each individual, including the individuals with nonrecurrent 8p rearrangements, are given in Table S1. Since individuals with invdupdel(8p) had similar deletion breakpoints

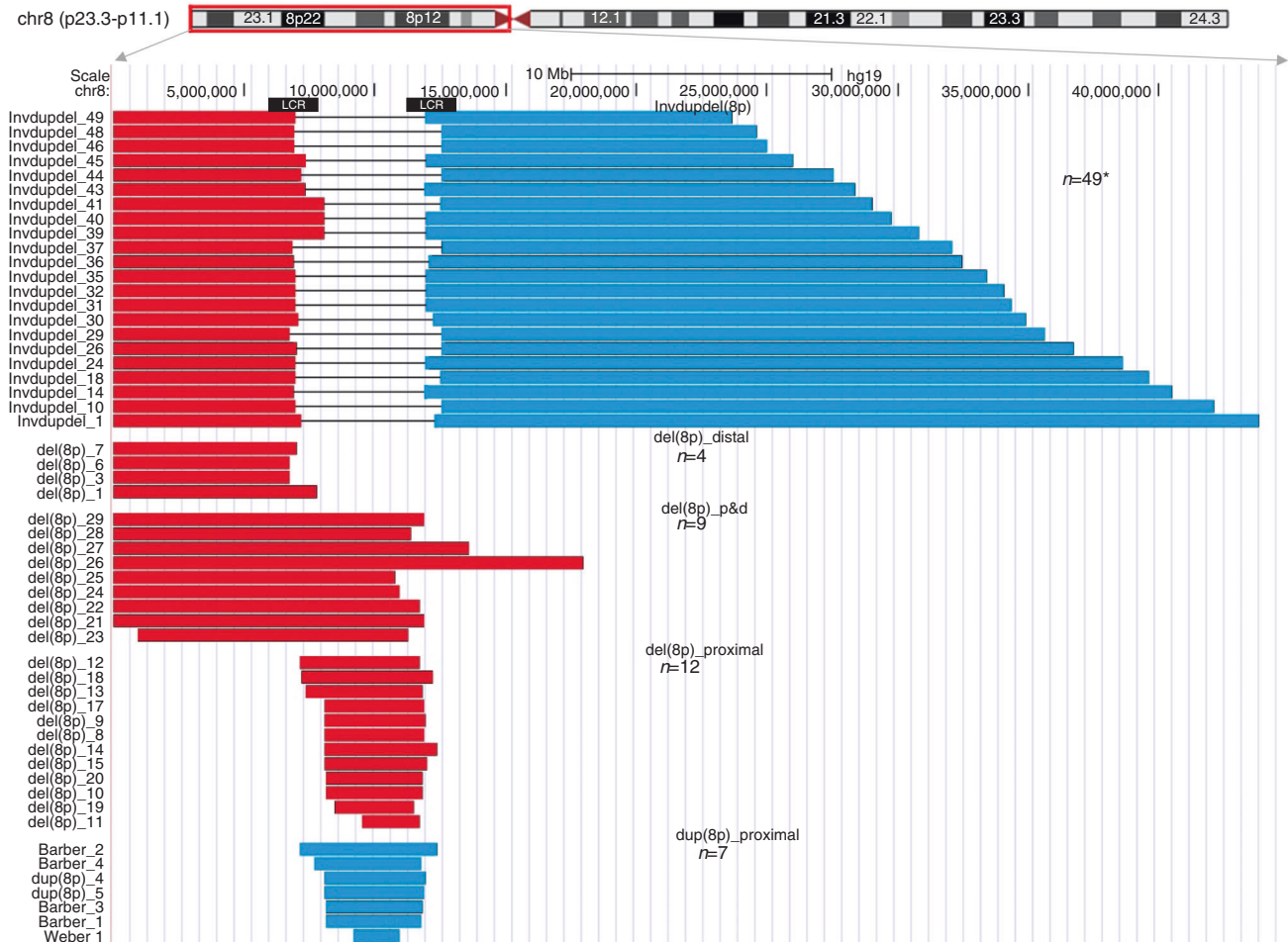


Fig. 1 A schematic representation of recurrent chromosome 8p rearrangements on University of California–Santa Cruz (UCSC) Genome Browser custom tracks. The ideogram of chromosome 8 is given at the top of the figure for scale and the shown area represents the entire short arm (p) of chromosome 8 as indicated by red rectangle. The solid black boxes below the coordinates ruler represent the low copy repeat (LCR) regions overlapping the breakpoint junctions. Red bars represent deletions, blue bars represent duplications. The thin black line between the red and blue bars at the top panel represent the copy-neutral interval between the deleted and duplicated segments of invdupdel(8p). *There are 22 individuals shown here but the tracks represent the rearrangements of all 49 individuals with invdupdel(8p) such that if two or more individuals had the identical rearrangement, only one of them was chosen to represent all individuals. For all individual tracks please refer to Figures S1, S2, and S3. p&d proximal & distal.

and sizes, and also similar proximal duplication breakpoints, they are grouped altogether given the similarity in clinical findings. However, seizures and corpus callosum abnormalities were separated by the distal breakpoint coordinates, and hence sizes, of their inverted duplications. We only report anthropometric measurements, developmental milestones, and cardiac findings for the dup(8p)_{proximal} cohort.

Prenatal and neonatal findings. Single umbilical artery, small for gestational age (SGA), and congenital heart defects are the most commonly reported findings in individuals with invdupdel(8p), del(8p)_{proximal&distal}, and del(8p)_{proximal} (Table 1). Oligohydramnios/polyhydramnios, intracranial cyst, reduced fetal movement, preeclampsia, and gestational diabetes mellitus were less frequently reported (Table S1).

The majority of individuals in each category had at least one neonatal finding with hypotonia being the most commonly reported. Respiratory distress, transient hypoglycemia and thermoregulation issues in term neonates, stiffness, torticollis, nuchal cord, polycythemia, and seizures were among other reported neonatal issues (Table 1 and Table S1).

The distributions of birth weight, length, and OFC Z-scores are shown in Figure S4, and the corresponding values in Table 1.

While most individuals across all rearrangements had Z-scores within 2 SD for all anthropometric measurements, individuals with del(8p)_{proximal} had lower average Z-scores for birth weight and length than individuals with other rearrangements. Individuals with del(8p)_{proximal} and del(8p)_{proximal&distal} had lower average Z-scores for birth OFC than individuals with other rearrangements. Of the individuals who were reported to be SGA prenatally, five of seven with invdupdel(8p), two of three with del(8p)_{proximal}, and three of three with del(8p)_{proximal&distal} were also born with birth weight less than the 10th percentile.

Growth and endocrine issues. Short stature was reported in 10–20% of individuals across all rearrangements (Table 1). Individuals with short stature generally had other issues such as past and/or current feeding difficulty along with lower concurrent weight Z-scores, history of being SGA, and/or prematurity (Table S1).

No individual with invdupdel(8p) was overweight or obese. However, 30–67% of individuals with del(8p) were overweight (Table 1), of whom only two adult individuals with del(8p)_{proximal&distal} and del(8p)_{distal} had short stature. Precocious puberty, hypothyroidism/

hyperthyroidism, diabetes mellitus, and low cortisol levels were among other occasionally reported endocrine issues (Table S1).

Neurodevelopmental milestones. The box plot distributions of developmental milestones (age at sitting, age at walking, and age at first words) are shown in Fig. 2. All individuals across all rearrangements had a delay in at least one developmental milestone with most individuals having global developmental delay. In general, individuals with invdupdel(8p) had more moderate-to-severe delays in all milestones compared to individuals with del(8p).

Of those who have spoken their first words, 4 of 15 individuals with invdupdel(8p) (average age = 48 months), 5 of 8 individuals with del(8p)_proximal (average age = 29.4 months), 5 of 7 individuals with del(8p)_proximal&distal (average age = 53 months), and all 3 individuals with del(8p)_distal (average age = 40.6 months) have also started speaking in short sentences. However, apart from four individuals with del(8p), all individuals have speech difficulties ranging from limited vocabulary to slurred speech, and many of them receive speech therapy. Speech difficulties are more severe in individuals with invdupdel(8p) than individuals with del(8p). The average age of 11 individuals with invdupdel(8p) who have spoken their first words but have yet to speak their first sentences was 70 months.

Although intelligence quotients were not available (many not testable), all individuals were reported to have some degree of intellectual disability and/or developmental delay (ID/DD), with individuals with del(8p) having milder neurocognitive deficits than individuals with invdupdel(8p). One adult with del(8p)_distal is attending university and another one was able to obtain a driver's license. The major issues in both individuals are mild learning disabilities and neurobehavioral/psychiatric issues such as short attention span, anxiety, and depression.

The average Vineland adaptive behavioral composite (ABC) scores were 52, 66, 45, and 66 for individuals with invdupdel(8p) ($n = 39$), del(8p)_distal ($n = 3$), del(8p)_proximal&distal ($n = 8$), and del(8p)_proximal ($n = 7$), respectively. Communication was the most severely affected subdomain in individuals with invdupdel(8p), del(8p)_distal, and del(8p)_proximal&distal (Figure S5).

Seizures. Twenty-seven individuals with invdupdel(8p), ten individuals with del(8p) (five with del(8p)_proximal, four with del(8p)_proximal&distal, and one with del(8p)_distal) were reported to have had at least one seizure (Table 1 and Table S1).

Seizure characteristics in individuals with del(8p). Each of the ten individuals across del(8p) subgroups had only one seizure type, and there was no prominent seizure type or recurrent seizure characteristics. In eight individuals, seizures started between 2 and 5 years of ages. Four of five individuals who had relatively more frequent seizures reported seizure control with sodium valproate. Five individuals have been followed without any antiepileptic medication.

Seizure characteristics in individuals with invdupdel(8p). Absence seizures were the most commonly reported seizure type in individuals with invdupdel(8p) followed by motor and febrile seizures (Table 1). Eleven of 27 individuals were reported to have more than one seizure type (Table S1).

The seizures usually started between 1 and 4 years of age ($n = 20$). The number of lifetime seizures tended to be either 1–10 or >100. Among the latter group, the frequency of seizures ranged between several times a day to 1–3 times a month. Absence seizures usually lasted less than a minute, and only two individuals with motor seizures reported one status epilepticus episode each.

There was no specific EEG abnormality in 36 individuals with invdupdel(8p) who had at least one EEG evaluation, yet half were

found to have an abnormal EEG (Table 1). Generalized slowing ($n = 9$, 26%) and epileptiform charges ($n = 7$, 20%) were the most commonly reported EEG abnormalities (Table S1). Four individuals were found to have nonspecific abnormal EEG without clinical seizures.

While there was no observable relationship between the occurrence, number, or frequency of absence seizures and distal duplication breakpoints, motor seizures were more commonly reported in individuals with more distal duplication breakpoints. While 9 of 25 individuals (36%) with distal duplication breakpoints after base pairs 37,000,000 had motor seizures, only 4 of 25 individuals (17%) with distal duplication breakpoints before base pairs 37,000,000 had motor seizures (Table S1).

Many individuals with invdupdel(8p) had medically manageable seizures. Levetiracetam, sodium valproate, oxcarbazepine, phenobarbital, topiramate, and phenytoin were used as monotherapy or in combinations. Twenty individuals no longer require antiepileptics.

Structural brain abnormalities. Structural brain abnormalities were commonly reported in individuals with invdupdel(8p), with agenesis/hypoplasia of the corpus callosum being the most common finding followed by hydrocephalus/ventriculomegaly and cerebral/cerebellar atrophy (Table 1). Most hydrocephalus cases were mild/benign and did not require shunt placement. Individuals with del(8p) occasionally had abnormal brain MRIs, and only one with del(8p)_distal was reported to have hypoplasia of the corpus callosum.

Of the individuals with distal duplication breakpoints before base pairs 37,000,000 ($n = 22/24$ with available brain MRIs), eight (36%) and one (5%) had hypoplasia and agenesis of the corpus callosum, respectively. Furthermore, only 2 of 12 individuals with the distal duplication breakpoints before base pairs 32,000,000 had hypoplasia of the corpus callosum. Of the individuals whose distal duplication breakpoints were after base pairs 37,000,000 ($n = 23/25$ with available brain MRIs), 8 (35%) had hypoplasia and 13 (57%) had agenesis of the corpus callosum.

Other neurological issues. Most individuals across all 8p subgroups had at least one neurological finding (Table S1). Generalized or truncal hypotonia was the most common neurological finding in all subgroups; however, it was more frequent and severe in individuals with invdupdel(8p) (Table 1 and S1). Hypotonia was accompanied by hypertonia, prominently in the lower extremities, later in life. Microcephaly was common in individuals with del(8p)_proximal. Only two individuals each with invdupdel(8p) and del(8p)_proximal&distal had microcephaly. Macrocephaly was reported in five individuals with invdupdel(8p); however, three of them had hydrocephalus/ventriculomegaly on brain MRI. Among the other less commonly reported neurological issues were developmental regression, gait difficulties, and tremors (Table S1).

Neurobehavioral issues. Neurobehavioral issues were among the major concerns in individuals with del(8p) compared to individuals with invdupdel(8p) (Table 1). Stereotypic behaviors and attention problems (attention deficit disorder [ADD], attention deficit–hyperactivity disorder [ADHD], hyperactivity) were the most commonly reported neurobehavioral issues in individuals with invdupdel(8p) and del(8p), respectively. Autism and autism spectrum traits, aggressivity/tantrums/impulsivity, sensory processing issues, echolalia, and anxiety were also commonly reported neurobehavioral issues (Table S1).

Congenital heart defects. Deleterious variants in *GATA4* on 8p23.1 are a known cause of congenital heart defects (CHD), particularly atrial septal defect (ASD) (MIM 607941), ventricular septal defect (VSD) (MIM 614429), atrioventricular septal defect

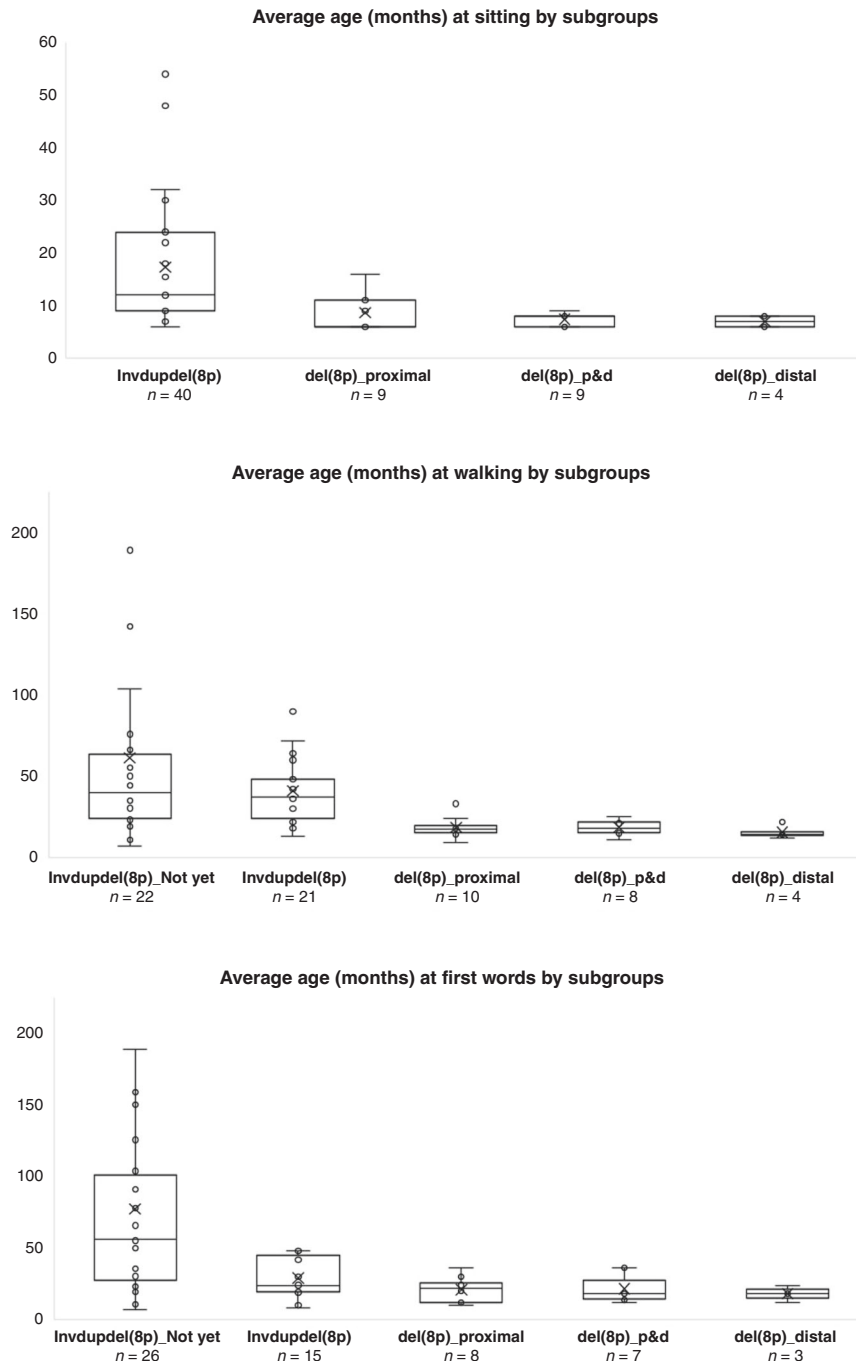


Fig. 2 The box plot distributions of average ages (months) of developmental milestones of sitting (top), walking (middle), and first words (bottom). All individuals across all rearrangements had a delay in at least one developmental milestone with most individuals having global developmental delay. Individuals with invdupdel(8p) had more moderate-to-severe delays in all milestones compared to individuals with del(8p), and over 50% of individuals with invdupdel(8p) have not achieved walking and/or speaking first words at the time of reporting (Invdupdel[8p]_Not yet). X represents the mean, the horizontal line within the boxes represents the median, and the lower and upper boundaries of the boxes represent the 25th and 75th quartiles, respectively. The number of individuals represented in each subgroup are given below their respective labels. p&d proximal & distal.

(ASVD) (MIM 614430), and tetralogy of Fallot (ToF) (MIM 187500). All individuals with deletions encompassing *GATA4* were found to have an ASD and/or VSD (Table S1). In addition to the septal defects, pulmonic stenosis (PS) was also reported in 75% of individuals with *GATA4* deletions. None of the individuals with VSD and PS in del(8p) subgroups was reported to have other components of ToF (misplaced/overriding aorta and right ventricular hypertrophy).

In recent years, *SOX7* (within del[8p]_proximal interval) has also been proposed to be associated with CHD and congenital diaphragmatic hernia (CDH) [12–14]. While deletions of 19 individuals with del(8p)_proximal and del(8p)_proximal&distal encompass both *GATA4* and *SOX7*, in two individuals with del(8p)_proximal&distal only *SOX7* was deleted, and only one of these individuals who was born at term had a small patent ductus arteriosus. Additionally, both genes were duplicated in two

individuals with reciprocal dup(8p)_proximal, and they were reported to have bicuspid aortic valve only. Of the 29 individuals whose deletions ($n = 21$) or duplications ($n = 8$) encompass SOX7, only 1 individual with del(8p)_proximal had CDH.

Both GATA4 and SOX7 were copy-number neutral in all individuals with invdupdel(8p) or del(8p)_distal. However, VSD and/or ASD were reported in over one third of individuals with invdupdel(8p) ($n = 15$, 38%), while no individual with del(8p)_distal was reported to have VSD and/or ASD. Additionally, patent foramen ovale (PFO) was also reported in seven individuals with invdupdel(8p) (18%), while no individual with del(8p) was reported to have PFO (Table S1).

In terms of CHD severity, structural heart defects tend to be minor and close spontaneously in individuals with invdupdel(8p) ($n = 19/26 = 73%$) compared to individuals with del(8p)_proximal and del(8p)_proximal&distal ($n = 4/16 = 25%$ in total) (data not shown). Of the four individuals with dup(8p)_proximal, two individuals from our cohort and two individuals from the literature were not reported to have had surgical repair. Among other less commonly reported cardiac defects across all individuals were dextraposition of great arteries, double outlet right ventricle, hypoplastic right heart, coarctation of aorta, aortic stenosis, aortic regurgitation, ascending aorta dilation, mitral valve prolapse, enlarged heart, and arrhythmia (Table S1).

Other organ systems and facial features. Other organ system abnormalities such as visual, gastrointestinal, musculoskeletal, genitourinary, immunological, dental, and sleep issues were also variably reported across all individuals (Table 1 and Table S1).

Figure 3 shows facial images of 11 individuals with various chromosome 8p rearrangements. Dysmorphic facial features in individuals with invdupdel(8p) include smaller head circumference, large and prominent forehead, mildly arched eyebrows, deep-set eyes, ptosis or hooded eyelids, full cheeks, wide mouth, and micrognathia (Fig. 3c, d, e). In individuals with del(8p)_proximal&distal (Fig. 3g, h, i), there are no distinctive facial features although rounded face, large forehead, full cheeks, low hanging columella, and minor dental issues can be appreciated. Individuals with del(8p)_proximal and dup(8p)_proximal do not have major dysmorphic facial features.

DISCUSSION

Delineation and comparison of the clinical spectrum of individuals who have deletion/duplication syndromes with different breakpoints is helpful to inform prognosis, tailor medical surveillance, and identify candidate genes that might contribute to the underlying clinical problems. In this study, we report a comprehensive clinical and cytogenomic summary of a large cohort of individuals with chromosome 8p rearrangements.

The most common clinical findings in individuals with chromosome 8p rearrangements have previously been reported in smaller less well characterized cohorts [15–17]. Owing to our large cohort size we were able to compare the frequency and severity of clinical findings across different 8p rearrangements (Table S1). The neurodevelopmental characteristics such as intellectual disability, seizures, tone abnormalities, and hypoplasia/agenesis of the corpus callosum are more frequent and/or more severe in individuals with invdupdel(8p) compared to individuals with del(8p); and the severity of these problems correlates with the breakpoints of the duplicated segment (Table 1 and S1). In contrast, individuals with del(8p) report more neurobehavioral issues.

Our results support GATA4 being the primary driver of CHD in individuals with del(8p)_proximal and del(8p)_proximal&distal. A third of individuals with invdupdel(8p) without GATA4 deletions also had VSD, ASD, and/or PS. Variants in MYOM2 located within the del(8p)_distal region have been reported with conotruncal

heart defects; however, our clinical data suggest other candidate gene(s) within the duplicated segment of invdupdel(8p) may contribute to CHD.

Individuals with chromosome 8p rearrangements share common nonspecific clinical findings associated with other neurodevelopmental conditions including feeding difficulties, constipation, and recurrent mild infections. A few individuals in our cohort had less common clinical findings such as choanal atresia, craniosynostosis, or tracheoesophageal fistula. To our knowledge, none of these individuals had exome or genome sequencing, and there could be other genetic factors contributing to these rare structural anomalies.

Given the large region of interest (entire chromosome 8p), identifying the smallest overlapping regions and candidate genes has been challenging, and only few genes on 8p have been assessed for disease evidence by ClinGen. Detailed discussion of candidate genes within each chromosome 8p rearrangement interval can be found in Supplemental Data. The large number of individuals allowed us to delineate the 11.0–11.5 Mb shared duplicated segment among individuals with invdupdel(8p). Three genes (TUSC3, VPS37A, RHOTB2) within and ten genes (NKX2-6, NEFL, CHRNA2, FZD3, TTI2, ERLIN2, PLPBP, ASH2L, DDHD2, KAT6A) beyond the shared duplicated segment have been associated with disorders characterized by overlapping neurodevelopmental phenotypes as seen in individuals with invdupdel(8p); however, with different inheritance modes and/or disease mechanisms (i.e. biallelic, loss-of-function, gain-of-function) (Supplemental Data and Table S2). Whether these genes are also triplosensitive and may cause the clinical findings reported in individuals with invdupdel(8p) is unknown.

Although the sizes of chromosome 8p rearrangements are above the classical microdeletion/microduplication syndromes (2–3 Mb), the same mechanism underlies the recurrent deletion/duplications of chromosome 8p. NAHR is mediated by LCR regions spanning base pairs 6,700,000–8,000,000 and 11,500,000–12,500,000 that are enriched for olfactory receptor and defensin gene clusters [4, 18, 19]. In the case of invdupdel(8p), a dicentric chromosome composed of deleted and inverted duplicated 8p segments may be formed following NAHR, and an ensuing random break within the inverted duplicated segment of 8p results in variable distal breakpoints [4]. This variability of the distal breakpoints further confounds the identification of additional candidate genes beyond shared duplicated segment given that the same major clinical findings are reported in individuals with the smallest and largest duplications although the severity differs with the distal duplication breakpoints. Thus, in addition to the gene(s) within the shared duplicated segment in invdupdel(8p), involvement of additional genes beyond the shared duplicated segment such as CHRNA2 may add incrementally to the clinical spectrum.

Identification of individuals with smaller deletions/duplications within the recurrent deletion/duplication intervals has facilitated identifying candidate genes for many deletion/duplication syndromes. We report three individuals with nonrecurrent deletions encompassing only DLGAP2 ($n = 1$) and CSMD1 ($n = 2$), whose neurodevelopmental problems are similar to the individuals with del(8p)_distal. Two of these deletions (DLGAP2 and CSMD1) were inherited, and inheritance of one deletion (CSMD1) is unknown in our cohort. Both genes have previously been proposed as candidates for neurodevelopmental and neurobehavioral phenotypes in individuals with del(8p) [15]. Animal models and functional studies support the roles of these genes in neurogenesis [20–24]. Although variants in these genes have been reported in individuals with autism spectrum disorders and schizophrenia, there is no established gene–disease association yet for either DLGAP2 or CSMD1 [25–27]. Future clinical and functional studies investigating these genes' role in neurodevelopment may be helpful to delineate their role in 8p.

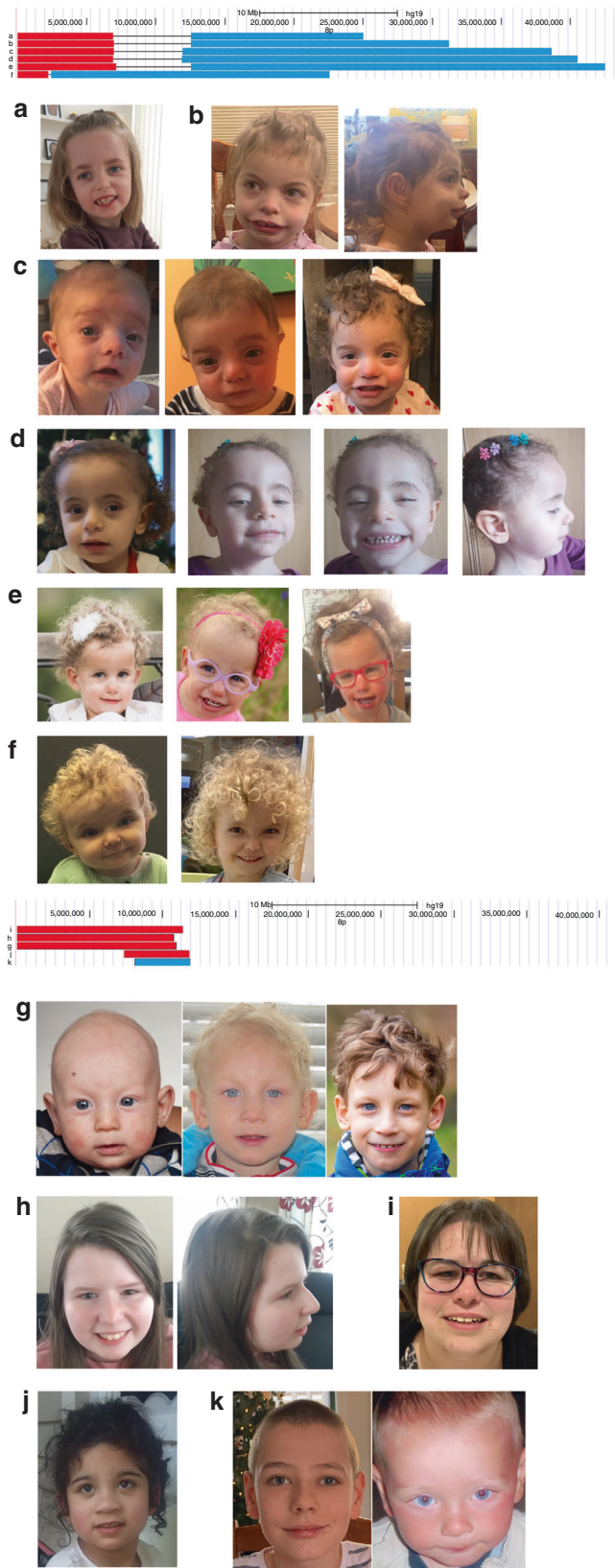


Fig. 3 Facial features of individuals with 8p cytogenomic disorders. Facial images of individuals with invdupdel(8p) (a–f), del(8p)_{proximal&distal} (g–i), del(8p)_{proximal} (j), and dup(8p)_{proximal} (k) with representative UCSC Genome Browser custom track of each individual's chromosomal rearrangement is given at the top with corresponding image letters. Dysmorphic facial features in individuals with invdupdel(8p) include smaller head with or without microcephaly, large and prominent forehead, mildly arched eyebrows, deep-set eyes, ptosis or hooded eyelids, full cheeks, wide mouth, and micrognathia (c–e). In individuals with del(8p)_{proximal&distal} (g–i), there are no distinctive facial features although rounded face, large forehead, full cheeks, low hanging columella, and minor dental issues can be appreciated. Individuals with del(8p)_{proximal} and dup(8p)_{proximal} do not have major dysmorphic facial features.

(8p) is lower than what would be expected based on the presumptive carrier frequency, and none of the families reported recurrence of any type of rearrangement in other children or recurrent miscarriages. There is also not a high frequency of invdupdel(8p) in molecular cytogenetic studies of products of conception. Thus, the absolute risk to 8p inversion carriers appears to be low.

Our study is limited by its reliance on caregiver report of clinical findings although we verified diagnoses with medical record review when possible. We also previously showed the validity of Vineland Adaptive Behavior Scales for assessing the developmental outcomes in individuals with congenital disorders [28].

In conclusion, we report detailed clinical and cytogenomic summary of 97 individuals with distinct chromosome 8p rearrangements with a summary of candidate genes. The clinical findings may be helpful to inform families and health-care providers about the associated features and their severity. The candidate genes may guide the researchers to investigate the underlying molecular and biological mechanisms of these disorders.

DATA AVAILABILITY

The data used in this study are available in Table S1.

Received: 27 November 2020; Revised: 29 June 2021; Accepted: 29 June 2021;

Published online: 19 July 2021

REFERENCES

- Carvalho CMB, Lupski JR. Mechanisms underlying structural variant formation in genomic disorders. *Nat Rev Genet.* 2016;17:224–38.
- Weckselblatt B, Rudd MK. Human structural variation: mechanisms of chromosome rearrangements. *Trends Genet.* 2015;31:587–99.
- Weleber RG, Verma RS, Kimberling WJ, Fieger HG, Lubs HA. Duplication-deficiency of the short arm of chromosome 8 following artificial insemination. *Ann Genet.* 1976;19:241–7.
- Giglio S, Broman KW, Matsumoto N, Calvari V, Gimelli G, Neumann T, et al. Olfactory receptor-gene clusters, genomic-inversion polymorphisms, and common chromosome rearrangements. *Am J Hum Genet.* 2001;68:874–83.
- Puvabanditsin S, Gengel N, Botti C, Jacob M, Jalil M, Cabrera K, et al. 8p 11 Microduplication is associated with neonatal stridor. *Mol Syndromol.* 2019;9:324–7.
- Barber JC, Rosenfeld JA, Foulds N, Laird S, Bateman MS, Thomas NS, et al. 8p23.1 duplication syndrome; common, confirmed, and novel features in six further patients. *Am J Med Genet A.* 2013;161A:487–500.
- Weber A, Köhler A, Hahn A, Müller U. 8p23.1 duplication syndrome: narrowing of critical interval to 1.80 Mbp. *Mol Cytogenet.* 2014;7:94.
- Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alfoldi J, Wang Q, et al. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature.* 2020;581:434–43.

It has been demonstrated that the invdupdel(8p) rearrangements occur on the maternal allele in 8p inversion carriers. It is estimated that 25% of individuals with European ancestry are carriers of this inversion [4]. However, the prevalence of invdupdel

9. Huang N, Lee I, Marcotte EM, Hurler ME. Characterising and predicting haploinsufficiency in the human genome. *PLoS Genet.* 2010;6:e1001154.
10. Chou JH, Roumiantsev S, Singh R. Peditools electronic growth chart calculators: applications in clinical care research, and quality improvement. *J Med Internet Res.* 2020;22:e16204.
11. Fenton TR, Kim JH. A systematic review and meta-analysis to revise the Fenton growth chart for preterm infants. *BMC Pediatr.* 2013;13:59.
12. Wat MJ, Beck TF, Hernández-García A, Yu Z, Veenma D, García M, et al. Mouse model reveals the role of SOX7 in the development of congenital diaphragmatic hernia associated with recurrent deletions of 8p23.1. *Hum Mol Genet.* 2012;21:4115–25.
13. Doyle MJ, Magli A, Estharabadi N, Amundsen D, Mills LJ, Martin CM. Sox7 regulates lineage decisions in cardiovascular progenitor cells. *Stem Cells Dev.* 2019;28:1089–103.
14. Yu L, Hernan RR, Wynn J, Chung WK. The influence of genetics in congenital diaphragmatic hernia. *Semin Perinatol.* 2020;44:151169.
15. García-Santiago FA, Martínez-Glez V, Santos F, García-Miñaur S, Mansilla E, Meneses AG, et al. Analysis of invdupdel(8p) rearrangement: Clinical, cytogenetic and molecular characterization. *Am J Med Genet A.* 2015;167A:1018–25.
16. Fisch GS, Davis R, Youngblom J, Gregg J. Genotype-phenotype association studies of chromosome 8p inverted duplication deletion syndrome. *Behav Genet.* 2011;41:373–80.
17. Ballarati L, Cereda A, Caselli R, Selicorni A, Recalcati MP, Maitz S, et al. Genotype-phenotype correlations in a new case of 8p23.1 deletion and review of the literature. *Eur J Med Genet.* 2011;54:55–9.
18. Liu P, Carvalho CMB, Hastings PJ, Lupski JR. Mechanisms for recurrent and complex human genomic rearrangements. *Curr Opin Genet Dev.* 2012;22:211–220.
19. Lupski JR. Genomic disorders: structural features of the genome can lead to DNA rearrangements and human disease traits. *Trends Genet.* 1998;14:417–22.
20. Distler MG, Opal MD, Dulawa SC, Palmer AA. Assessment of behaviors modeling aspects of schizophrenia in Csm1 mutant mice. *PLoS One.* 2012;7:e51235.
21. Steen VM, Nepal C, Erslund KM, Holdhus R, Nævdal M, Ratvik SM, et al. Neuropsychological deficits in mice depleted of the schizophrenia susceptibility gene CSMD1. *PLoS One.* 2013;8:e79501.
22. Kraus DM, Elliott GS, Chute H, Horan T, Pfenninger KH, Sanford SD, et al. CSMD1 is a novel multiple domain complement-regulatory protein highly expressed in the central nervous system and epithelial tissues. *J Immunol.* 2006;176:4419–30.
23. Deneault E, White SH, Rodrigues DC, Ross PJ, Faheem M, Zaslavsky K, et al. Complete disruption of autism-susceptibility genes by gene editing predominantly reduces functional connectivity of isogenic human neurons. *Stem Cell Rep.* 2018;11:1211–25.
24. Jiang-Xie L-F, Liao HM, Chen CH, Chen YT, Ho SY, Lu DH, et al. Autism-associated gene Dlgap2 mutant mice demonstrate exacerbated aggressive behaviors and orbitofrontal cortex deficits. *Mol Autism.* 2014;5:32.
25. Rasmussen AH, Rasmussen HB, Silahatoglu A. The DLGAP family: neuronal expression, function and role in brain disorders. *Mol Brain.* 2017;10:43.
26. Athanasiu L, Giddaluru S, Fernandes C, Christoforou A, Reinvang I, Lundervold AJ, et al. A genetic association study of CSMD1 and CSMD2 with cognitive function. *Brain Behav Immun.* 2017;61:209–16.
27. Schizophrenia Working Group of the Psychiatric Genomics Consortium. Biological insights from 108 schizophrenia-associated genetic loci. *Nature.* 2014;511:421–7.
28. Wynn J, Aspelund G, Zygumt A, Stolar CJ, Mychaliska G, Butcher J, et al. Developmental outcomes of children with congenital diaphragmatic hernia: a multicenter prospective study. *J Pediatr Surg.* 2013;48:1995–2004.

ACKNOWLEDGEMENTS

We thank all the individuals and their families for their participation, and specifically to Project 8p (www.project8p.org) and Unique (www.rarechromo.org). This study was partially supported by Project 8p and the JPB Foundation.

AUTHOR CONTRIBUTIONS

Conceptualization: W.K.C. Formal analysis: V.O., H.K., W.K.C., B.L. Funding acquisition: W.K.C. Investigation: V.O., W.K.C. Methodology: V.O., W.K.C. Project administration: V.O., L.H., H.K., C.M., S.R. Writing—original draft: V.O., W.K.C. Writing—review & editing: V.O., L.H., H.K., C.M., S.R., B.L., W.K.C.

ETHICS DECLARATION

This study was approved by the institutional review board of Columbia University (IRB AAAA5719). We obtained written consent to use deidentified patient data from participants. An additional consent for publication of facial photos was obtained from all participants presented in Fig. 3.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41436-021-01270-2>.

Correspondence and requests for materials should be addressed to W.K.C.

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.