

# Low-level parental mosaicism affects the recurrence risk of holoprosencephaly

Ping Hu, MS<sup>1</sup>, Ariel F. Martinez, PhD<sup>1</sup>, Paul Kruszka, MD<sup>1</sup>, Seth Berger, MD, PhD<sup>1</sup>,  
Erich Roessler, MD, PhD<sup>1</sup> and Maximilian Muenke, MD<sup>1</sup>

**Purpose:** De novo variants (DNVs) represent an important fraction of the pathogenic variant burden in holoprosencephaly (HPE). However, unexpected recurrences can occur, as evidenced by multiple affected children harboring the same apparently DNV. This study was performed to estimate the rate of parental mosaicism in a cohort of patients with HPE.

**Methods:** We developed a targeted capture next-generation sequencing (NGS) panel of 153 genes with potential implication in HPE. Sequencing data from a cohort of 136 HPE family trios were analyzed to identify probands with apparently DNVs. DNVs were examined in the proband and their parents to detect any deviations from the expected ~50/50 allele ratio of true heterozygosity. Selected variants were confirmed by Droplet Digital™ polymerase chain reaction (ddPCR).

**Results:** We identified 28 high-confidence DNVs, 20 of which

occurred in known HPE genes. Nineteen of the 20 variants (95%) were pathogenic or likely pathogenic. Sequence data analysis showed evidence of parental mosaicism in five cases, for an overall mosaicism rate of 26%. In addition, we found evidence for likely postzygotic events in four cases (50%).

**Conclusions:** High sensitivity methods, such as high-depth NGS and ddPCR, are essential to providing an accurate assessment of recurrence risk in HPE families with apparently DNVs.

*Genetics in Medicine* (2019) 21:1015–1020; <https://doi.org/10.1038/s41436-018-0261-8>

**Keywords:** holoprosencephaly; somato-gonadal mosaicism; next-generation sequencing; droplet digital PCR; genetic counseling

## INTRODUCTION

Recent advances in next-generation sequencing (NGS) technologies have begun to uncover previously underreported disease mechanisms implicating somatic variation in cancer,<sup>1</sup> overgrowth syndromes,<sup>2</sup> and novel case reports of somato-gonadal mosaicism in genetic syndromes.<sup>3</sup> These NGS advances have improved our understanding of environmental cofactors, maternal factors, and parental age effects on mutation rates and spectra.<sup>4,5</sup>

Mosaicism refers to genomic variation that is detected in some tissues and not others. Postzygotic variants can be limited to a subset of organs or tissues and may include the germline stem cell populations (somato-gonadal mosaicism). Parental gonadal mosaicism for apparently de novo variants (DNVs) carries the additional risk of transmission to subsequent offspring (typically only 0.1% of all DNVs).<sup>5</sup> Although parental mosaicism has been reported in a variety of contexts, including copy-number and single-nucleotide variants, and for a number of conditions, including epilepsy, it has not been systemically investigated in most disorders, including holoprosencephaly (HPE).<sup>3,6,7</sup>

HPE is the most common structural malformation of the brain and face in humans occurring in 1:250 embryos, but only 1:10,000 live-born children.<sup>8</sup> Genes associated with HPE are under strong purifying selection and are often intolerant to deleterious variation. Most cases of HPE occur in children born to apparently healthy parents consistent with either sporadic genetic or environmental causes. A significant number of HPE cases are considered to be caused by DNVs.<sup>9</sup> Clear examples of gonadal mosaicism affecting multiple sibs are well documented, although rare. Studies of individuals within families who harbor pathogenic variants demonstrate both incomplete penetrance and variable expressivity, the cause of which remains largely unexplained.<sup>10</sup> Both males and females are equally affected and neither biased parental transmission nor effects of parental age have been established.<sup>9</sup> The epidemiological findings are consistent across all human populations.<sup>11</sup> Genes validated to be associated with HPE function during a brief window of a shared vertebrate developmental system that explains the common phenotypic spectrum and its potential overlap with other midline conditions.<sup>12</sup> Common known HPE genes (*ZIC2*, *SHH*,

<sup>1</sup>Medical Genetics Branch, National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland, USA. Correspondence: Maximilian Muenke ([mamuenke@mail.nih.gov](mailto:mamuenke@mail.nih.gov))

Submitted 7 February 2018; accepted: 26 July 2018

Published online: 10 September 2018

*SIX3*, and *FGFR1*) satisfy all of the pathogenicity criteria of autosomal dominant malformation syndromes, with little evidence for obligate modifiers.<sup>13</sup> Teratogens can cause midline signaling abnormalities with maximal effect during the same developmental window and impacting the same genetic programs.<sup>14</sup> We set out to perform a retrospective study of 136 father–mother–child trios with HPE to identify DNVs and determine the rate of parental mosaicism.

## MATERIALS AND METHODS

We analyzed 136 family trios with a child affected with HPE. All families provided informed consent for our genetic evaluations as monitored by the National Human Genome Research Institute Institutional Review Board (clinicaltrials.gov: NCT00088426). The coding regions of 153 developmental genes and their regional noncoding elements (putative enhancers, promoters, untranslated regions, intronic elements, etc.) visually demonstrating a high overall conservation (>70% over >100 bp) were selected for our targeted capture strategy, using the ECRbrowser (<https://ecrbrowser.dcode.org/>) and the University of California–Santa Cruz (UCSC) browser (<https://genome.ucsc.edu>). Approximately 1 Mb of sequence was interrogated per individual with an overall coverage of 97.5%. High-confidence DNVs were identified and classified by gene or gene locus. Raw data supporting the variant detections were further investigated in the Integrative Genomics Viewer (IGV, <http://www.broadinstitute.org/igv/>) using accepted criteria and quality control annotations. Confirmation of parental mosaicism in blood was performed by Droplet Digital™ polymerase chain reaction (ddPCR) using the QX200 system (Bio-Rad, Hercules, CA). ddPCR experiments were repeated three times.

## RESULTS

All 28 high-confidence DNVs detected are summarized in Table 1, Supplemental Table S1 and Supplemental Table S2. All parents in this cohort previously had negative test results by Sanger sequencing for the variants identified in their affected offspring. After removing low-quality findings and repeat regions from further consideration, we detected 28 DNVs, 20 of which (71%) occurred in known HPE genes: *ZIC2* (8 cases), *SIX3* (5 cases), *SHH* (4 cases), and *FGFR1* (3 cases). Following current American College of Medical Genetics and Genomics (ACMG) guidelines,<sup>15</sup> 19 of these 20 variants were classified as pathogenic or likely pathogenic (18 missense and 1 affecting a canonical splice site), and 1 as having uncertain significance. All 19 pathogenic or likely pathogenic variants were confirmed by Sanger sequencing in our CLIA lab or by GeneDx (genedx.com). Two of these variants (*SIX3* p.[E129\*] and *SHH* p.[E256\*]) have been previously reported in apparently healthy individuals in the Kaviar Genomic Variant database (<http://db.systemsbiology.net/kaviar>) with a single observation each.

Among the 19/20 pathogenic variants in known HPE genes, each of the four disease genes has at least one example of high

confidence somato-gonadal mosaicism in the parental blood samples combined with a consistent family history and physical exam (Fig. 1 and Table 1): family 1: *SIX3* p.(W113\*); family 2: *SHH* p.(E256\*); family 3: *FGFR1* p.(R627T); family 4: *ZIC2* p.(V326Afs\*88); and family 5: *ZIC2* c.1076-1 G>A. Family 1 had a previous affected child who succumbed to alobar HPE and was not available for testing. Families 2 and 3 have additional affected children carrying the same variant. The mother in family 4 has microform HPE and is a mosaic for the pathogenic variant identified in her child. We do not know whether family 5 had previous affected pregnancies because family history was incomplete. Therefore, at least 5/19 (26%) of these cases are best explained by mosaicism conferring an elevated risk of recurrence in subsequent pregnancies. As shown in Fig. 1, all five gene variants were confirmed by ddPCR.

Given the number of families included in our cohort and the size of our target region, we expected to find, by chance alone, seven DNVs in genes not likely related to HPE. Therefore, the additional eight variants (in non-HPE genes, see Supplemental Table S2) represent an increase in variation burden revealing new potential candidate genes. These occur in both coding regions (three findings: *SCUBE1* p.[G398E], *NKX2-2* p.[G26D], and *IFT27* p.[R138Q]), intronic regions (four findings, three of which are poorly conserved), and one poorly conserved 3' untranslated region (UTR) variant (*ACVR1B* c.\*2201G>A). The three coding variants in non-HPE genes have been observed in healthy individuals in gnomAD (<http://gnomad.broadinstitute.org>), although at a very low frequency. *NKX2-2* p.(G26D) was predicted to be damaging by multiple algorithms (CADD<sub>phred</sub> score = 24.2) and was classified as a variant of uncertain significance. Despite the fact that we explored a similar number of noncoding targets to exons (~500,000 bp in each group) there was a striking enrichment of de novo findings in the four well-established HPE genes (4/153 gene targets and 20/28 detections), compared with noncoding elements. Interestingly, for four of these findings in Supplemental Table S1 and Table S2 (see BL11527, LCL6463, BL5443, and BL9276) the alternative (A)-to-reference (R) allele ratio in the proband is unusually low (11–28%), suggesting a postzygotic de novo event.

We identified five additional variants in classical HPE genes that could not be confirmed by ddPCR (Supplemental Table S1): *SHH* p.(S112\*), *SHH* p.(Y435N), *ZIC2* p.(H286Rfs\*80), *ZIC2* p.(H286Q), and *SIX3* p.(H155P). Given the lower quality metrics of the targeted capture data these findings likely represent false positives.

## DISCUSSION

Few studies have addressed mosaicism and its potential clinical impact. Some cohort studies have reported somatic mosaicism in genes associated with epilepsy<sup>3,6</sup> and overgrowth syndromes,<sup>2</sup> but for most genes the information is limited to case reports. While a proportion of mosaic individuals have disease manifestations, others are silent carriers of pathogenic variants that go undetected until they

**Table 1** Summary of targeted capture BAM files analysis of parental mosaicism in families with putative de novo variants detected in probands and confirmation analysis by ddPCR

Proband <sup>a</sup>	Variant	Read count in proband [% = A/(A + R)] {ACMG interpretation} <sup>a</sup>	Read count in parents F [% = A/(A + R)] M [% = A/(A + R)]	Mosaicism confirmation by ddPCR (%)	Clinical findings: proband {relative(s) and/or controls}
Family 1 BL9306 <sup>b</sup>	<i>SIX3</i> c.338G>A <sup>c</sup> p.(W113*)	[A = 51% = 229/445] {pathogenic}	F [0% = 0/361] M [9% = 37/406]	F (not tested) M (13%)	Lobar HPE, corpus callosum dysgenesis, hydrocephalus, spina bifida, seizures, developmental delay. {Deceased sib with alobar HPE, not tested for this variant.}
Family 2 BL3969	<i>SHH</i> c.766G>T p.(E256*)	[A = 46% = 31/68] {pathogenic}	F [10% = 10/99] M [0% = 0/85]	F (13%) M (not tested)	Proband has HPE. {Sib AM3808 confirmed as affected and c.766G>T, p.(E256*) positive by CLIA testing at GeneDx. One detection among healthy controls in Kaviar.}
Family 3 BL9658 <sup>b</sup>	<i>FGFR1</i> c.1880G>C p.(R627T)	[A = 47% = 340/724] {pathogenic}	F [3% = 26/840] M [0.1% = 1/827]	F (3.7%) M (not tested)	Semilobar HPE, corpus callosum dysgenesis, diabetes insipidus, growth hormone deficiency, bilateral cleft lip/palate, microtia, cryptorchidism, ectrodactyly both hands and feet. {Multiple affected sibs testing positive for this variant. <sup>9</sup> }
Family 4 BL11025 <sup>b</sup>	<i>ZIC2</i> c.975_976dupCG p.(V326Afs*88)	[A = 32% = 581/1828] {pathogenic}	F [0%, 0/1291] M [1.54% = 19/1233]	F (not tested) M (2.1%)	Normal karyotype, alobar HPE, hypotelorism, depressed nasal root. {Mother has microform findings, microcephaly, hypotelorism, learning disability.}
Family 5 BL7158 <sup>b</sup>	<i>ZIC2</i> c.1076- 1G>A	[A = 49% = 76/155] {pathogenic, splice site}	F [0.3% = 1/342] M [0% = 0/195]	F (0.1%) M (not tested)	Semilobar HPE. Incomplete family history.

A alternative allele; ACMG American College of Medical Genetics and Genomics; BL blood sample; ddPCR Droplet Digital™ polymerase chain reaction; F father; HPE holoprosencephaly; LCI lymphoblastoid cell line; M mother; R reference allele

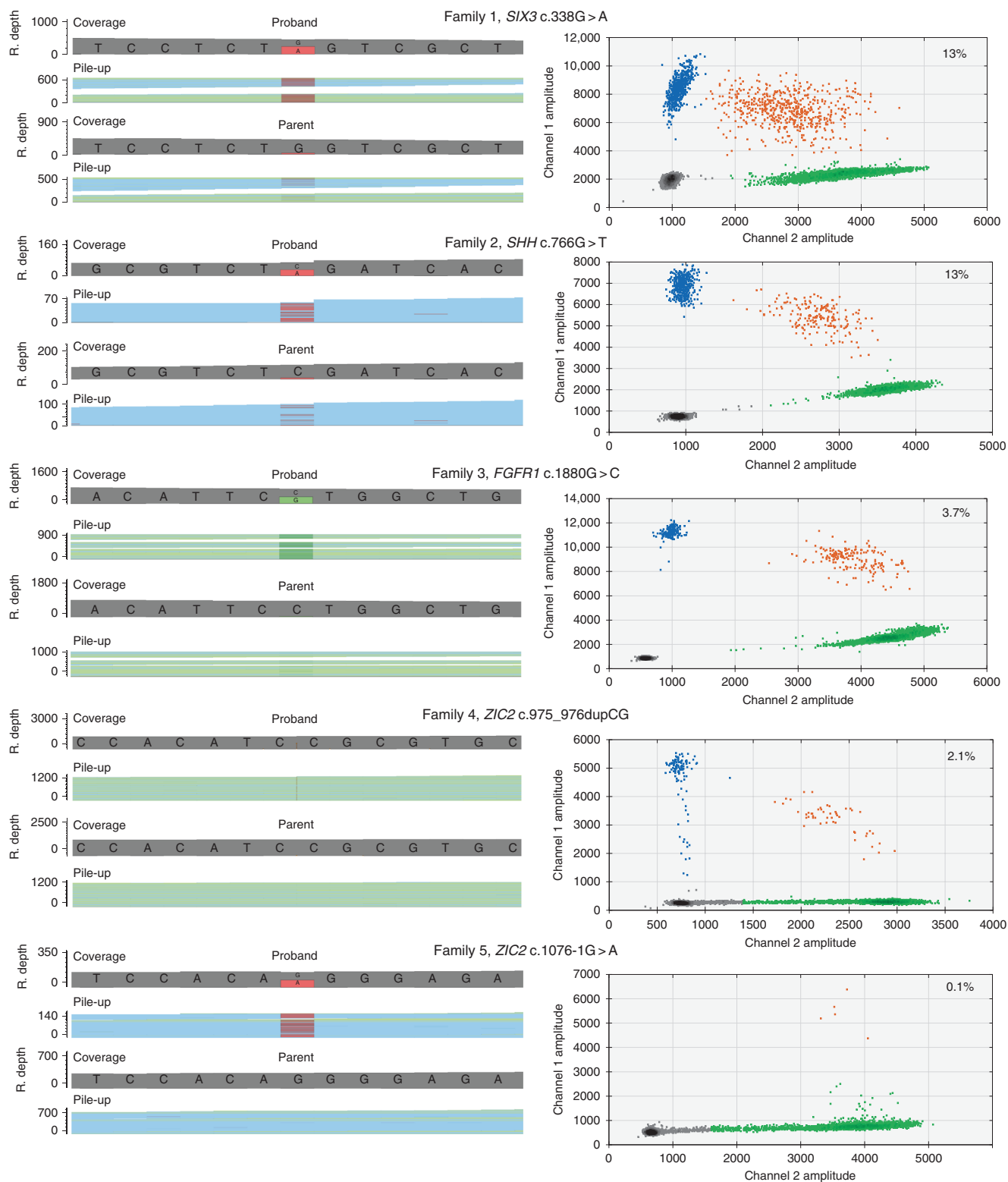
<sup>a</sup>Richards et al.<sup>15</sup>

<sup>b</sup>Confirmed by Sanger sequencing in a CLIA lab

<sup>c</sup>variants correspond to these transcripts: NM\_000193.2 (*SHH*), NM\_007129.3 (*ZIC2*), NM\_005413.3 (*SIX3*), NM\_023110.2 (*FGFR1*)

<sup>d</sup>Hong et al. *Hum Mol Genet* 2016;25(10):1912–1922

# BRIEF COMMUNICATION



surface in affected offspring. Low-level mosaicism is usually missed by technologies such as Sanger and low read depth NGS (e.g., exome sequencing); thus, a higher than expected proportion of children with variants regarded as *de novo* may be born to parents with somato-gonadal mosaicism. This has

obvious implications for the estimation of recurrence risk in affected families.

Our present work provides evidence for somato-gonadal mosaicism affecting genes that have been well replicated for nearly two decades of HPE clinical and molecular studies. We

**Fig. 1** Families with parental somato-gonadal mosaicism identified by next-generation sequencing and confirmed by Droplet Digital™ polymerase chain reaction (ddPCR). For each family, read depth information from the BAM files (left) indicates that the variant/reference allele ratio was ~50%/50% in the proband, but not in the carrier parent. ddPCR analysis (right) confirmed the variant allele was in fact underrepresented in the parent. 2D plots show the FAM fluorescence amplitude (channel 1, variant allele) and HEX fluorescence amplitude (channel 2, reference allele) for each droplet. Percentages of variant-positive versus variant-negative droplets were used to calculate the level of mosaicism (top right corner of each plot) using the Poisson statistics, as implemented in the QuantaSoft™ software (Bio-Rad, Hercules, CA). Blue dots, droplets containing variant alleles only; green dots, droplets containing reference alleles only; red dots, droplets containing both alleles; black dots, droplets containing no alleles. The lower limit of detection was defined as the presence of at least two positive droplets in three independent experiments.

found that this is the case in a minimum of 26% of families with pathogenic or likely pathogenic variants (5/19 variants). This was documented clinically in three of the families, and both clinically and molecularly in two of them. This mosaicism rate is similar to that found by Yang *et al.*,<sup>6</sup> who identified mosaicism in 25% of parents among 112 families with Dravet syndrome (due to *SCN1A* variation). In another large study, Myers *et al.*<sup>3</sup> detected parental mosaicism in 8.3% of families with Dravet syndrome or other epileptic encephalopathies. While the rate can vary for different genetic conditions, these results suggest that parental mosaicism may be more common than previously thought.

As noted above, only four of our *de novo* findings were consistent with a postzygotic pathogenic mechanism in the proband. It is interesting to speculate that given the narrow window of HPE pathogenesis (the third week of gestation, prior to the separation of the soma and gonadal progenitors) the absence of a strong phenotype in pathogenic variant-positive mosaic parents reflects the actual timing of occurrence of the DNV. A nearly 50% variant allele ratio in proband tissues would indicate either an inherited allele, or an extremely early postzygotic event involving tissues that participate in brain development.

The lowest level of mosaicism detected by ddPCR in our study was  $10^{-3}$  (0.1%, family 5, Fig. 1), which is higher than the  $10^{-4}$  detection limit previously reported in cancer and epilepsy samples.<sup>1,6</sup> Our cases of parental mosaicism detected by targeted capture sequencing and ddPCR could not be detected by Sanger sequencing, which has been the gold standard for diagnostic laboratories for several decades. This highlights the importance of using more sensitive technologies in clinical genetic testing. Additional validation using ultrasensitive NGS approaches, such as duplex sequencing<sup>16</sup> or o2n-seq,<sup>17</sup> may help to confirm very low levels of mosaicism measured by ddPCR at different sensitivity levels.

DNVs in five families could not be confirmed by ddPCR (Supplemental Table S1). Although these variants were undetectable in the parents' peripheral blood, they could be present in other tissues including the germline. Previous publications have reported changes in the variation spectrum and rate in parental germline cells,<sup>5</sup> and they stress the importance of including paternal sperm samples in genetic testing. Germline variants detected in the fathers of probands affected by diseases such as Noonan syndrome (caused by *PTPN11* variants), Apert syndrome (caused by *FGFR2* variants), and Costello syndrome (caused by *HRAS* variants),

have been previously studied.<sup>18–20</sup> Those studies show evidence of an accumulation of mosaic variants and an elevation of variant allele frequency in germline cells. Given the risk of complications with invasive prenatal testing, prenatal diagnosis is not routinely suggested for a second pregnancy following the birth of a child with a simplex case of a genomic disorder. However, if screening for parental somatic mosaicism is able to identify families with an increased recurrence risk, prenatal diagnosis might be offered. Such prospective analyses would have certainly changed recurrence risk counseling for families 1–5, where the evidence of mosaicism could have affected choice or management of subsequent pregnancies. Therefore, detection and confirmation of low-level parental mosaicism must be offered to all at-risk families.

Another important aspect of mosaicism analysis involves the quality of the databases with which the patient/family data are compared. Are the exomes of 1 million people better than 10,000 genomes done at greater depth and quality? Other questions may arise: does somatic mosaicism help to explain those rare pathogenic findings among healthy cohorts? To what extent does somatic mosaicism predict a risk for transmission to offspring? The answer to these questions will shape the future of DNA sequencing technologies and provide the tools for a more accurate genetic risk assessment.

Replication of our results will require the commitment to a novel strategy for variant detection in the future. We recommend the use of more sensitive technologies, the routine testing of paternal sperm samples, and the analysis of multiple parental peripheral tissues from different embryonic origins (e.g., ectoderm: buccal swab; mesoderm: blood; and endoderm: urine).

#### ELECTRONIC SUPPLEMENTARY MATERIAL

The online version of this article (<https://doi.org/10.1038/s41436-018-0261-8>) contains supplementary material, which is available to authorized users.

#### ACKNOWLEDGEMENTS

The authors wish to thank the patients and their health-care providers who participated in this research. This research was supported by the Division of Intramural Research (DIR), National Human Genome Research Institute, the National Institutes of Health. We thank the National Intramural Sequencing Center (NISC) for sponsoring the Flagship series of intramural clinical/molecular studies and their participation in the design and implementation of this targeted capture study and whose full

contributions are presented as a separate manuscript (Roessler et al., in preparation).

## DISCLOSURE

The authors declare no conflicts of interest.

## REFERENCES

1. Watanabe M, Kawaguchi T, Isa S, et al. Ultra-sensitive detection of the pretreatment EGFR T790M mutation in non-small cell lung cancer patients with an EGFR-activating mutation using Droplet Digital PCR. *Clin Cancer Res*. 2015;21:3552–3560.
2. Chang F, Liu L, Fang E, et al. Molecular diagnosis of mosaic overgrowth syndromes using a custom-designed next-generation sequencing panel. *J Mol Diagn*. 2017;19:613–624.
3. Myers CT, Hollingsworth G, Muir AM, et al. Parental mosaicism in “de novo” epileptic encephalopathies. *N Engl J Med*. 2018;378:1646–1648.
4. Biesecker LG, Spinner NB. A genomic view of mosaicism and human disease. *Nat Rev Genet*. 2013;14:307–320.
5. Rahbari R, Wuster A, Lindsay SJ, et al. Timing, rates and spectra of human germline mutation. *Nat Genet*. 2016;48:126–133.
6. Yang X, Liu A, Xu X, et al. Genomic mosaicism in paternal sperm and multiple parental tissues in a Dravet syndrome cohort. *Sci Rep*. 2017;7:15677.
7. Rothlisberger B, Kotzot D. Recurrence risk in de novo structural chromosomal rearrangements. *Am J Med Genet A*. 2007;143A:1708–1714.
8. Shiota K, Yamada S. Early pathogenesis of holoprosencephaly. *Am J Med Genet*. 2010;154C:22–28.
9. Solomon BD, Bear KA, Wyllie A, et al. Genotypic and phenotypic analysis of 396 individuals with mutations in Sonic Hedgehog. *J Med Genet*. 2012;49:473–479.
10. Solomon BD, Mercier S, Vélez JJ, et al. Analysis of genotype-phenotype correlations in human holoprosencephaly. *Am J Med Genet*. 2010;154C:133–141.
11. Orioli IM, Castilla EE. Epidemiology of holoprosencephaly: prevalence and risk factors. *Am J Med Genet*. 2010;154C:13–21.
12. Hong S, Hu P, Roessler E, Hu T, Muenke M. Loss-of-function mutations in FGF8 can be independent risk factors for holoprosencephaly. *Hum Mol Genet*. 2018;27:1989–1998.
13. Roessler E, Hu P, Muenke M. Holoprosencephaly in the genomics era. *Am J Med Genet C Semin Med Genet*. 2018. <https://doi.org/10.1002/ajmg.c.31615> PMID:29770992
14. Hong M, Krauss RS. Ethanol itself is a holoprosencephaly-inducing teratogen. *PLOS ONE*. 2017;12:e0176440.
15. Richards S, Aziz N, Bale S, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med*. 2015;17:405–423.
16. Kennedy SR, Schmitt MW, Fox EJ, et al. Detecting ultralow-frequency mutations by duplex sequencing. *Nat Protoc*. 2014;9:2586–2606.
17. Wang K, Lai S, Yang X, et al. Ultrasensitive and high-efficiency screen of de novo low-frequency mutations by o2n-seq. *Nat Commun*. 2017;8:15335.
18. Giannoulatou E, McVean G, Taylor IB, et al. Contributions of intrinsic mutation rate and selfish selection to levels of de novo HRAS mutations in the paternal germline. *Proc Natl Acad Sci USA*. 2013;110:20152–20157.
19. Goriely A, Wilkie AO. Paternal age effect mutations and selfish spermatogonial selection: causes and consequences for human disease. *Am J Hum Genet*. 2012;90:175–200.
20. Yoon SR, Choi SK, Eboreime J, Gelb BD, Calabrese P, Arnheim N. Age-dependent germline mosaicism of the most common noonan syndrome mutation shows the signature of germline selection. *Am J Hum Genet*. 2013;92:917–926.