Germline mutation in POLR2A: a heterogeneous, multi-systemic developmental disorder characterized by transcriptional dysregulation

Adam W. Hansen,1,2,* Payal Arora,3 Michael M. Khayat,1,2 Leah J. Smith,4 Andrea M. Lewis,2 Linda Z. Rossetti,2 Joy Jayaseelan,1 Ingrid Cristian,5 Devon Haynes,5 Stephanie DiTroia,6 Naomi Meeks,7 Mauricio R. Delgado,8,9 Jill A. Rosenfeld,2,10 Lynn Pais,8 Susan M. White,11,12 Qingchang Meng,1 Davut Pehlivan,2,13 Pengfei Liu,2,10 Marie-Claude Gingras,1,2,14 Michael F. Wangler,2,15 Donna M. Muzny,1,2 James R. Lupski,1,2,16,17 Craig D. Kaplan,3,* and Richard A. Gibbs1,2

Summary

De novo germline variation in POLR2A was recently reported to associate with a neurodevelopmental disorder. We report twelve individuals harboring putatively pathogenic de novo or inherited variants in POLR2A, detail their phenotypes, and map all known variants to the domain structure of POLR2A and crystal structure of RNA polymerase II. Affected individuals were ascertained from a local data lake, pediatric genetics clinic, and an online community of families of affected individuals. These include six affected by de novo missense variants (including one previously reported individual), four clinical laboratory samples affected by missense variation with unknown inheritance—with yeast functional assays further supporting altered function—one affected by a de novo in-frame deletion, and one affected by a C-terminal frameshift variant inherited from a largely asymptomatic mother. Recurrently observed phenotypes include ataxia, joint hypermobility, short stature, skin abnormalities, congenital cardiac abnormalities, immune system abnormalities, hip dysplasia, and short Achilles tendons. We report a significantly higher occurrence of epilepsy (8/12, 66.7%) than previously reported (3/15, 20%) (p value = 0.014196; chi-square test) and a lower occurrence of hypotonia (8/12, 66.7%) than previously reported (14/15, 93.3%) (p value = 0.076309). POLR2A-related developmental disorders likely represent a spectrum of related, multi-systemic developmental disorders, driven by distinct mechanisms, converging at a single locus.

Introduction

The human enzyme DNA-directed RNA polymerase II (EC 2.7.7.6) transcribes all nuclearly encoded messenger RNA (mRNA). It is a large enzyme composed of twelve subunits, the largest of which—the 220-kDa subunit A—is encoded by POLR2A (MIM: 180660). POLR2A contains essential domains of the RNA polymerase II enzyme, including the catalytic core and a C-terminal heptapeptide repeat, the differential phosphorylation of which is critical for regulating transcriptional dynamics.1,2 Numerous structural and mutational studies in various systems have been conducted, revealing a spectrum of genetic variants differentially impacting distinct dimensions of transcription (i.e., initiation, elongation, etc.).3–4 Indeed, the function of RNA polymerase II has been extensively studied for decades.

Despite its centrality within the central dogma of molecular biology and its extensive study over decades, POLR2A was not implicated in human disease until 2016, when Clark et al.5 reported multiple distinct, recurrent somatic mutations in the gene as causative for a clinically unique subset of meningiomas. Very recently the first report of pathogenic germline mutations in POLR2A was published, describing a phenotypically heterogeneous neurodevelopmental syndrome with hypotonia (MIM: 618603).6 Here, we report additional clinical and molecular evidence strengthening the case for POLR2A dysfunction as a multi-systemic, phenotypically heterogeneous Mendelian disorder.

Material and methods

DNA sequencing and genotyping

For individuals 1, 6, and 8–13, DNA capture and sequencing of exomes was carried out as previously described by Hansen et al.7 at either the Baylor Genetics (BG) laboratories or at the Baylor

1Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX, USA; 2Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX, USA; 3Department of Biological Sciences, University of Pittsburgh, Pittsburgh, PA, USA; 4Department of Biochemistry and Biophysics, Texas A&M University, TX, USA; 5Division of Genetics, Arnold Palmer Hospital for Children, Orlando Health, Orlando, FL, USA; 6Broad Center for Mendelian Genomics and Program in Medical and Population Genetics, Broad Institute of MIT and Harvard, Cambridge, MA, USA; 7Departments of Pediatrics and Genetics, University of Colorado School of Medicine, Aurora, CO, USA; 8Texas Scottish Rite Hospital for Children, Dallas, TX, USA; 9Department of Neurology, University of Texas Southwestern Medical Center, Dallas, TX, USA; 10Baylor Genetics Laboratories, Houston, TX, USA; 11Vicorian Clinical Genetics Services, Murdoch Children’s Research Institute, Parkville 3052, VIC, Australia; 12Department of Paediatrics, University of Melbourne, Parkville, VIC, Australia; 13Section of Pediatric Neurology and Developmental Neuroscience, Department of Pediatrics, Baylor College of Medicine, Houston, TX, USA; 14Michael E. DeBakey Department of Surgery, Baylor College of Medicine, Houston, TX, USA; 15Jan and Dan Duncan Neurological Research Institute, Texas Children’s Hospital, Houston, TX, USA; 16Texas Children’s Hospital, Houston, TX, USA; 17Department of Pediatrics, Baylor College of Medicine, Houston, TX, USA
*Correspondence: awphd@gmail.com (A.W.H.), craig.kaplan@pitt.edu (C.D.K.)
https://doi.org/10.1016/j.xhgg.2020.100014.
© 2020 The Authors. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
College of Medicine Human Genome Sequencing Center (HGSC). Sequencing and analysis for individuals 2 (genome sequencing) and 5 (exome sequencing [ES]) were provided by the Broad Institute of MIT and Harvard Center for Mendelian Genomics (Broad CMG). ES and analysis for individuals 3, 4, and 7 was performed at other commercial clinical laboratories. Chromosomal microarray analysis (CMA) for individual 8 was performed at BG. CMA for individuals 1–8, 10, and 12 were performed at other commercial clinical laboratories. It is unknown whether CMA was performed for individuals 9 and 11.

NGS analysis
Initially, a local data lake containing ES data for approximately 20,000 individuals with suspected Mendelian disorders (Hadoop Architecture LakE of Exomes [HARLEE]) was utilized to discover POLR2A as a candidate Mendelian disease-associating gene. Within this dataset, fastq files were aligned to hg19, and variants were called with Atlas2 (v1.4.3) and annotated with VEP. High-quality ultra-rare (MAF < 1/10,000) variants observed in individuals within HARLEE were prioritized. This analysis resulted in the discovery of ultra-rare, potentially pathogenic POLR2A variants in individuals 1, 6, and 8–13. Individuals 2–5 and 7 were ascertained for this study after the initial published discovery of a POLR2A-related developmental disorder, with analysis conducted by the Broad CMG or other commercial clinical laboratories.6

For all individuals, region-specific intolerance to missense variants is calculated with the Missense Tolerance Ratio (MTR) score,6 with scores < 1.0 indicating a lower-than-expected ratio of missense to synonymous variants in the ExAC dataset1 for the 31-bp window surrounding an amino acid residue. Estimates of residue-level conservation were obtained from GERP++.10 via the UCSC Genome Browser.11

Phenotyping
Phenotyping is described in the Supplemental material and methods, with all phenotypes summarized in Table 1.

Structural domains and alignment
POLR2A structural domains were derived from the yeast RNA polymerase II (Pol II) crystal structure, which shares a remarkably high level of conservation with POLR2A.3 Human POLR2A domain coordinates were derived by alignment as performed with MAFFT FFT-NS-2 (v7.305b).12

Functional evaluation of variants in yeast
Thirteen ultra-rare variants in POLR2A were identified in clinical samples at the time the yeast experiments were initiated. Yeast studies were conducted in the yeast ortholog of POLR2A, RPO21 (generally referred to as RPB1), and are summarized in Table 2. Detailed experimental methods are described in the Supplemental material and methods.

Ethics statement
Data for individuals 1–8 were collected after written informed consent in conjunction with the Baylor Hopkins Center for Mendelian Genomics (CMG) (H-29697) study with approval by the institutional review board at Baylor College of Medicine. Other clinical samples (individuals 9–12) were from the Baylor College of Medicine clinical testing laboratories, now incorporated as BG; these data were studied in aggregate for the purpose of improving the diagnostic assay, under protocol H-41191. The procedures followed were in accordance with the ethical standards of the responsible committee on human experimentation (institutional and national), and proper informed consent was obtained.

Results
Utilizing the HARLEE data lake,7 a total of seven clinical exomes were originally identified with ultra-rare (MAF < 1/10,000) suspected pathogenic variants in POLR2A. Attempts were made to contact all individuals and recruit to a research protocol (see Material and methods). After the initial report of de novo variation in POLR2A causing a Mendelian disorder,6 five additional individuals were subsequently identified through the genetics clinic at Texas Children’s Hospital and a social media-based support group for families of individuals diagnosed with pathogenic variants in POLR2A (Table 1), including one previously published individual for whom amended and additional phenotypic information is provided (individual 7; Haijes et al.,6 individual 15). The family of individual 7 reports that she is positive for a few phenotypes previously reported as negative: feeding difficulty/failure to thrive, decreased endurance, and decreased fetal movement. Phenotypic data from clinical samples detected in HARLEE for individuals without research consent are reported in aggregate, in accordance with institutional review board (IRB)-approved protocols (see Ethics statement).

Clinical ES or genome sequencing revealed no other genetic diagnosis for any of these individuals. CMA reportedly did not reveal any findings for individuals 1–8 and 12. It is unknown whether CMA was performed for individuals 9 and 11. CMA for individual 10 reportedly identified a variant of uncertain significance (VUS) gain in Xp22.13. Thus, in total, twelve individuals with ultra-rare, putatively pathogenic variants in POLR2A are reported. Pathogenicity for all of these variants is supported by the extreme degree of constraint for missense (missense intolerance Z score = 8.59) and loss-of-function (pLI = 1.0) variants observed in healthy individuals from the ExAC dataset.9 All variants occur at highly conserved residues (as indicated by a GERP++ score ≥ 2.0), in regions further constrained for missense variation in the ExAC dataset (Table 1).8,10 Of these twelve variants, there are ten affected by missense variants (including one previously published individual and two individuals affected by the same variant). We also report one individual affected by an in-frame deletion and one family affected by a C-terminal frameshift variant. Variants are distributed throughout the length of the protein product, with no obvious association between severity of phenotype and affected protein domain (Figure 1). We report one variant located within the Clamp core domain, two variants within the Clamp head, one variant within the Dock, three variants within the Cleft, one variant within the Trigger loop, one variant within the Jaw (observed in two individuals), and one
Table 1. Variant information and phenotypes of individuals with ultra-rare POLR2A variants

| Individual ID | 1     | 2     | 3     | 4     | 5     | 6     | 7 (Haijz et al.; individual 15) | 8     | 9     | 10    | 11    | 12    | Summary |
|---------------|-------|-------|-------|-------|-------|-------|------------------------------|-------|-------|-------|-------|-------|---------|
| Consent to sharing of individual-level phenotypes? | yes   | yes   | yes   | yes   | yes   | yes   | yes                          | yes   | no    | no    | no    | no    | NA      |
| Sex           | F     | M     | M     | M     | F     | F     | F                           | F     | M     | F     | M     | M     | NA      |
| Age (in years, at requisition or most recent phenotyping) | 0.32  | 4     | 3.5   | 10    | 21    | 6.75  | 11                          | 14    | 1.13  | 16.12 | 1.49  | 4.36  | NA      |
| Molecular | | | | | | | | | | | | |
| Previous/ additional Molecular Diagnosis | unsolved | unsolved | unsolved | unsolved | unsolved | unsolved | mitochondrial complex IV deficiency indicated via skeletal muscle enzyme analysis; mtDNA sequencing was negative | unsolved | unsolved | unsolved | unsolved | unsolved | NA      |
| Chromosome | 17    | 17    | 17    | 17    | 17    | 17    | 17                          | 17    | 17    | 17    | 17    | 17    | NA      |
| Position | 7401503 | 7411610 | 7411736 | 7412890 | 7412890 | 7415280 | 7416391                     | 7417020 | 7388166 | 7399625 | 7399813 | 7411604 | NA      |
| Genomic variant | GACCCTIC>G  | C>T    | C>T   | A>G   | A>G   | G>A   | G>A                         | CCA>C  | C>G   | G>A   | C>T   | C>T   | NA      |
| cDNA variant | c.1314_1319del | c.3281C>T | c.3407C>T | c.3752A>G | c.4252G>T | c.4808G>A | c.5440_5441del            | c.83C>G | c.323G>A | c.418C>T | c.3275C>T | NA      |
| Amino acid change | p.Leu438_His439del | p.Ser1094Phe | p.Thr1113Ile | p.Asm1251Ser | p.Asm1251Ser | p.Gly1418Arg | p.Arg1603His         | p.Gln1814Val699ter | p.Pro28Arg | p.Arg108His | p.Arg1407P | p.Ala1092Val | NA      |
| GERP++ score (conservation) | mean = 4.64 (0.89–5.6; SD = 1.9; 6 bases) | 5.36  | 5.23  | 5.11  | 5.11  | 5.06  | 3.93                         | mean = 1.68 (-8.25–4.13; SD = 3.06; 476 bases) | 5.82  | 5.57  | 5.57  | 5.36  | NA      |
| MTR score (31 bp) (missense constraint) | mean = 0.675 (0.637–0.678; 2 amino acids) | 0.477 | 0.172 | 0.516 | 0.516 | 0.301 | 1.031                        | mean = 0.7 (0.448–0.968; SD = 0.151; 158 amino acids) | 0.598 | 0.649 | 0.673 | 0.528 | NA      |
| Zygosity | het    | het   | het   | het   | het   | het   | het                         | het    | het   | het   | het   | het   | NA      |
| De novo in proband? | yes   | yes   | yes   | yes   | yes   | yes   | yes                         | no; inherited from mother with mild difficulty learning | ?     | ?     | ?     | ?     | NA      |
| Yeast variant phenotype | transcription defect (weak) (yeast Ile424D corresponds to human p.Leu438del) | ?     | ?     | ?     | ?     | ?     | ?                           | protein defect (strong) | ?     | ?     | ?     | ?     | NA      |

(Continued on next page)
| Individual ID | 1 | 2 | 3 | 4 | 5 | 6 | 7 (Haijes et al.\(^\d\) individual 15) | 8 | 9 | 10 | 11 | 12 | Summary |
|---------------|---|---|---|---|---|---|-------------------------------------|---|---|----|----|----|---------|
| **Phenotypes** |   |   |   |   |   |   |                                     |   |   |    |    |    |         |
| Short stature (HP:0004322) | + | – | – | – | – | – | 5th percentile | – | – |    |    |    | 3/4     |
| **Head and neck** |   |   |   |   |   |   |                                     |   |   |    |    |    |         |
| Dysmorphic features (HP:0001999) | + |   |   |   |   |   | prominent supraorbital ridges; high posterior hairline; upslanted palpebral fissures; fair complexion; deeply set eyes; strabismus (when tired) | – |   |   |   |   |   | NP |
| Abnormal brain MRI (HP:0012443) | polymicrogyria (HP:0002126), ventriculomegaly (HP:0002119), hydrocephalus (HP:0000228), hypomyelination (HP:0006808) | Rathke cleft cyst | – |   |   |   | prominent third and lateral ventricles at 1 year old (HP:0007082, HP:0006956) | – |   |   |   |   |   | 1/4 |
| Microcephaly/ macrocephaly | NP | – | – | – | – | – | significant microcephaly (<2 percentile) | – | – | – | – | – | macrocephaly 1/4 | 3/12 |
| Visual impairment (HP:0000505) | NP | – | – | – | – | – | bilateral elongation of orbital globes noted on MRI | – | – | – | – | – | – | NP |
| Hearing impairment (HP:0000365) | NP | – | – | – | – | – | sensorineural | – | – | – | – | – | right-side hearing deficit attributed to hemiplegia | NP |
| Headache (HP:0002315) | NP | – | – | – | – | – | reported by parents; cannot confirm as individual is non-verbal | – | – | – | – | – | daily severe headaches | NP |

(Continued on next page)
| Individual ID | 1 | 2 | 3 | 4 | 5 | 6 | 7 (Haijes et al. 2006) | 8 | 9 | 10 | 11 | 12 | Summary |
|---------------|---|---|---|---|---|---|-----------------------|---|---|----|----|----|---------|
| **Central nervous system** |   |   |   |   |   |   |                       |   |   |    |    |    |         |
| Developmental delay (HP:0012758) | + | + | + | + | + | + | speech and motor (HP:0000750, HP:0001270) | + | + | speech and motor (HP:0000750, HP:0001270) | 4/4 | 12/12 |
| Developmental regression (HP:0002376) | NP | three separate 6-month episodes of regression and plateau | – | – | – | – | regression at age 16 years attributed to worsening anxiety, with new aggressive behaviors | – | – | no regression, but developed personality changes (HP:0000751) and memory problems (HP:0002354) at age 8 years | NP | 4/12 |
| Seizures (HP:0001250) | NP | EEG revealed right centroparietal epileptiform discharges | – | – | – | – | generalized tonic-clonic and absence during childhood, with no discernable seizures since puberty, no longer on anticonvulsants | + | + | started at 13 years old | 2/4 | 8/12 |
| Intellectual disability (HP:0001249) | NP | + | + | + | + | + | + | mild | 1/4 | 8/12 |
| Autism spectrum disorder (HP:0000729) | NP | + | + | + | + | + | + | – | – | 1/4 | 6/12 |
| Hypotonia (HP:0001252) | – | – | – | – | – | – | – | – | – | 1/4 | 8/12 |
| Dystonia (HP:0001332) | NP | – | + | + | – | – | – | – | – | hypertonia/ spasticity of the right extremities due to hemiplegia (HP:0001276, HP:0004374) | NP | 3/12 |
| Neck and back arching in infancy | NP | + | + | + | – | – | NP | – | – | NP | – | 3/12 |
| Abnormal movements (HP:0100022) | NP | hand flapping, kicking, rubs side of head | head banging | occasional hand flapping | hand flapping | head hitting,stimming arms and legs | – | – | – | 1/4 | 7/12 |
| Pica (HP:0001856) | NP | eats hair, paper, feces | frequently puts objects in mouth | NP | NP | chews on hands, toes, clothes, paper, furniture, toys | NP | NP | NP | 3/12 |
| Ataxia (HP:0001251) | NP | + | + | + | + | + | + | ataxic high-stepping gait | + | + | ataxic gait | NP | 7/12 |

(Continued on next page)
| Individual ID | 1 | 2 | 3 | 4 | 5 | 6 | 7 (Hajjes et al.) | 8 | 9 | 10 | 11 | 12 | Summary |
|---------------|---|---|---|---|---|---|------------------|---|---|----|----|----|---------|
| Difficulty sleeping<br>(HP:0002360) | NP | obstructive sleep apnea prior to removal of tonsils and adenoids; sleeps >12 h and exhibits daytime sleepiness | – | + | +<br>woke up 10x per night in childhood, current symptoms are occasional and mild | NP | no sleep study, but reportedly very light sleeper | insomnia, takes clonidine for it | NP | 5/12 |
| Skeletal abnormality<br>(HP:0000924) | NP | – | – | – | – | – | mild pectus excavatum (HP:0000767), hip dysplasia (HP:0001385), sixth lumbar vertebrae (HP:0008416) | NP | hip dysplasia (HP:0001385), pes planus (HP:0001763), leg length discrepancy (HP:000559) | contractures on right side (HP:0001371), two Achilles tendon release surgeries (HP:0001771) | 1/4 | 5/12 |
| Scoliosis<br>(HP:00002650) | NP | – | – | – | – | – | neurogenic | – | – | – | NP | 1/12 |
| Joint hypermobility<br>(HP:0001382) | NP | + | + | + | + | – | hyperextensible elbows and knees | + | – | – | 1/4 | 6/12 |
| Failure to thrive<br>(HP:0001508)/feeding difficulty<br>(HP:00011968) | + | + | – | + | – | – | poor weight gain but normal linear growth | – | – | – | 2/4 | 7/12 |
| Uses gastrostomy tube | NP | planned but not yet inserted | – | inserted at 9 years old | – | inserted at 5 years old | – | – | – | – | NP | 3/12 |
| GI reflux<br>(HP:0002020) | NP | + | – | + | + | – | – | – | – | 1/4 | 5/12 |
| Constipation<br>(HP:0002019)/diarrhea<br>(HP:0002014) | NP | constipation | constipation | constipation | – | – | constipation and diarrhea | – | – | – | diarrhea 1/4 | 5/12 |
| Urogenital system<br>(HP:0000119) | NP | incontinence | incontinence | neurogenic bladder; incontinence; swollen left kidney after birth; small scrotum, retractile testes | toilet trained at age 6 years, still some incontinence | NP | incontinence | history of several urinary tract infections (HP:0000010) | kidney abnormality 1/4 | 7/12 |
| Skin abnormality<br>(HP:0000951) | NP | easy scarring, keratosis pilaris | – | easy scarring | easy scarring | keratosis pilaris, facial flushing | – | – | – | unspecified abnormality 1/4 | 5/12 |

(Continued on next page)
| Individual ID | 1 | 2 | 3 | 4 | 5 | 6 | 7 (Hajesi et al.) individual 15 | 8 | 9 | 10 | 11 | 12 | Summary |
|--------------|---|---|---|---|---|---|-----------------------------|---|---|---|---|---|---------|
| Immune system (HP:0002715) | NP | neutropenia (HP:0001875); granulocytopenia (HP:0001913); high lymphocyte percentage (HP:0100827) | – | intermittent leukocytosis (HP:0001974) | – (no antibiotics needed for >10 years) | NP | – | – | – | unspecified abnormality 1/4 | 3/12 |
| Fever (HP:0001954) | NP | chronic (HP:0001955) | – | fever of unknown etiology (HP:0001955) | fever of unknown etiology in childhood (HP:0001955) | – | – | – | – | chronic fever 1/4 | 4/12 |
| Cardiac abnormality (HP:0001627) | dilated cardiomyopathy (HP:0001644) | – | Bicuspid aortic valve (HP:0001647) | Atrophic septal defect (HP:0001631), congestive heart failure in infancy (HP:0001635) | – | NP | – | – | – | NP | 3/12 |
| Lung abnormality (HP:0002086) | immature lungs (HP:0006703) | recurrent upper respiratory infections (HP:0002788); frequent nebulizer use (budesonide and albuterol) | recurrent upper respiratory infections (HP:0002788); uses inhaler and nebulizer | poor airway clearance requiring vest therapy, suctioning, and cough assist; recurrent upper respiratory infections (HP:0002788) | – | – | – | recurrent upper respiratory infections (HP:0002788); hospitalized four times for pneumonia, once for bronchiolitis | NP | 5/12 |
| Congenital diaphragmatic hernia (HP:0000776) | NP | – | – | – | – | – | – | – | – | – | 1/4 | 1/12 |
| Other | NP | hypothyroidism (HP:0000821); tooth misalignment (HP:0006692); consistently elevated BUN/creatinine ratio (HP:0004081); alkaline phosphatase (HP:0004379), and prolactin (HP:0004086) | – | history of angioedema with levetiracetam use (HP:0100665); hemangioma (HP:0001028); 2-vessel umbilical cord (HP:0001195); adverse reaction to therapeutic botulinum toxin; enlarged liver that resolved after birth (HP:0006564); hyponatremia (HP:0002902) | psychosis induced by fluoxetine; agitation with sedatives and general anesthesia | NP | intolent of wheat and dairy (not allergic); reduced fetal movement (HP:0001558) | – | NP | 4/12 |

All variants are considered likely pathogenic. Phenotypes are indicated as either positive, negative, or not phenotyped or reported (NP) for each individual. For research-consented individuals, individual-level phenotypic information is presented. For the four other clinical individuals (individuals 9–12), phenotypic information is presented in aggregate. The summary column indicates how many individuals were positive for a given phenotype. All cDNA variants are based on the NM_000937.4 transcript definition. All genomic coordinates are based on the hg19 genome build. Human Phenotype Ontology (HPO) IDs are indicated for phenotypes where appropriate. +, positive; −, negative; BUN, blood urea nitrogen; GI, gastrointestinal; NA, not applicable.
frameshift variant in the C-terminal domain (CTD), which is likely to escape nonsense-mediated decay as it occurs within the last exon of the gene.

**Missense variants**

Of the ten individuals harboring missense variants, four non-research-consented individuals have phenotypes reported in aggregate (individuals 9–12) rather than at an individual level. Six of these ten missense variants were confirmed to be inherited de novo. The remaining four were observed in clinical samples for which parental ES was not performed, but their pathogenicity is further supported by functional evidence and phenotypic similarity. Commonly reported phenotypes among these individuals include developmental delay (7/7), seizures (≥7/10), hypotonia (≥7/10), abnormal movements (≥7/10), ataxia (≥6/10), autism spectrum disorder (ASD) (≥6/7), failure to thrive/feeding difficulty (≥6/10), joint hypermobility (≥6/10), abnormal brain MRI (≥5/10), incontinence (≥5/10), skin abnormalities including keratosis pilaris and easy scarring (≥5/10), visual impairment (≥4/10), short stature (≥4/10), difficulty sleeping (≥4/10), skeletal abnormalities (≥4/10), recurrent upper respiratory infections (≥4/10), and cool distal extremities (≥4/10). (Diagnoses of intellectual disability and autism spectrum disorder cannot be ruled out for three clinical samples, as pheno-typing for these individuals is limited to what was included on the ES requisition, ordered for these individuals at an age too young to typically diagnose intellectual disability or ASD.)

Individuals 4 and 5 are of note, sharing an identical variant with a previously reported individual: c.3752A>G (p.Asn1251Ser; Haijes et al.6) Individual 14 was reported as a 6-year-old girl with hypotonia, strabismus, frog position in infancy, decreased endurance, feeding difficulties, recurrent respiratory tract infections, disturbed sleeping, gastro-esophageal reflux, failure to thrive, microcephaly, brachyplagiocephaly, decreased fetal movements, aggressive behavior, pectus excavatum, walking at 5.5 years of age, and mega cisterna magna.6 A detailed clinical description of individuals 4 and 5 is included in the supplement of this manuscript (Supplementary note). There is considerable variability in age of walking across these three individuals, reported as 3.5 (individual 5), 4.5 (individual 4), and 5.5 (Haijes et al.,6 individual 14) years of age. Other notable phenotypic differences include the presence of a cardiac abnormality.
atrial septal defect) in individual 4 and the presence of recurrent respiratory infections in individual 4 and Haijes et al. individual 14, but not in individual 5. Facial dysmorphology is relatively similar for individuals 4 and 5 (Figure 2).

Other variants
We report one frameshift variant that is potentially pathogenic—g.7417023_7417024del (GenBank: NC_000017.10) (c.5440_5441del [NM_000937.4]; p.Gln1814Valfs99ter)—observed in individual 8, with a remarkably...
more mild presentation than the individuals affected by missense variants. This variant occurs in the last exon of the gene (29/29) and leads to a premature termination near the C-terminal end of the protein, after amino acid residue 1912/1970, and an alteration of 99 amino acids. This alteration would result in truncation of the CTD by 20 heptapeptide repeats and 1 partial out of 52. Truncation of CTD repeats can confer phenotypes in model organisms.\textsuperscript{14–20} Furthermore, given its position in the gene, it is predicted to escape nonsense-mediated decay, with a fully expressed, yet truncated, protein product. Individual 8 presented at 8 years of age with hypertonia/spasticity of the right extremities, recent worsening headache, memory problems and personality changes, hemiplegia, and delayed speech and motor milestones. MRI of the brain revealed left hippocampal atrophy, and an electroencephalogram (EEG) revealed a focus of spike activity in the left central region. Clinical ES initially failed to detect any pathogenic variants consistent with phenotypic presentation. At most recent follow-up, individual 8 was reported to have developed seizures at age 13 years, recurrent urinary tract infections, and a history of two Achilles tendon release surgeries. Daily severe headaches were reported, as well as difficulty sleeping. Aggregate genocentric reanalysis through HARLEE\textsuperscript{7} revealed individual 8 as positive for the ultra-rare C-terminal frameshift variant in \textit{POLR2A}. Sanger sequencing of proband and maternal saliva samples revealed the variant to be maternally inherited. The mother of individual 8 reports no significant related medical history other than delayed speech and mild learning difficulty. Maternal family history is also positive for a son who was born at 35 weeks and had hypoplastic left heart, brain anomaly, and failure to thrive. He had three open heart surgeries and died at 7 months of age. He was reported to have a chromosome 6p duplication with an unknown \textit{POLR2A} genotype.

Harboring an in-frame deletion of two amino acids, individual 1 presented with the most severe phenotype of all evaluated individuals. Clinical ES was ordered at 3 months of age, with reported phenotypes of immature lungs, dilated cardiomyopathy, failure to thrive, hypotonia, developmental delay, dysmorphic features, and abnormal brain MRI findings including polymicrogyria, ventriculomegaly, hydrocephalus, and hypomyelination. ES failed to detect any known pathogenic variants consistent with phenotypic presentation. Individual 1 reportedly died during infancy. As above, genocentric reanalysis identified an ultra-rare variant in \textit{POLR2A}—chr17:g.7401503GACCTTC>G (NM_000937.4) \textsuperscript{(c.1314_1319del; p.His439_Leu440del). The average GERP score across the six deleted nucleotides is 4.64, indicative of high evolutionary conservation. Notably, this same variant, when present as a developmental somatic mutation, has been established as causal for a subset of meningiomas.\textsuperscript{5} Meningioma was not reported in individual 1. Sanger sequencing of parental saliva samples failed to detect the variant. Sanger sequencing in individual 1 confirmed the presence of the heterozygous variant.
Functional studies

Functional assays were conducted in yeast to further evaluate the pathogenicity of observed POLR2A variants in clinical and research samples (Figure 3). As noted above, the large subunit of Pol II (encoded by POLR2A in human, RPB1 in yeast) is highly conserved in sequence and structure. We previously established a number of plate phenotypes highly predictive of transcription defects due to specific alterations to Pol II catalytic activity in yeast.21,22 Residues analogous to some identified in individuals had been identified previously as being mutated in genetic screens for yeast transcription mutants; rpb1 Pro24Ser was identified as rpb1-915 (analogous residue to POLR2A p.Pro28), while rpb1 Gly1388Val was identified as sua8-423 (analogous to POLR2A p.Gly1418). Here we employed these tests to interrogate conserved residues impacted by missense variants observed in humans for growth defects in yeast. The yeast strains utilized were the same as in Hajies et al.,6 as their strains were derived from the Kaplan lab. Mutant plasmids encoding variants in conserved residues identified in a subset of individuals were introduced into yeast as the sole copy of RPB1 and phenotyped on a number of growth media. We observed conditional growth defects as well as phenotypes related to altered transcription for a subset of mutants (Table 2). Conditional defects such as temperature sensitivity or formamide sensitivity are consistent with protein folding or assembly defects exacerbated by heat or solvent, and these were observed are consistent with protein folding or assembly defects such as temperature sensitivity or formamide sensitivity such as temperature sensitivity or formamide sensitivity such as temperature sensitivity or formamide sensitivity such as temperature sensitivity or formamide sensitivity such as temperature sensitivity or formamide sensitivity such as temperature sensitivity or formamide sensitivity such as temperature sensitivity or formamide sensitivity such as temperature sensitivity or formamide sensitivity such as temperature sensitivity or formamide sensitivity such as temperature sensitivity or formamide sensitivity such as temperature sensitivity or formamide sensitivity such as temperature sensitivity or formamide sensitivity such as temperature sensitivity or formamide sensitivity such as temperature sensitivity or formamide sensitivity such as temperature sensitivity or formamide sensitivity such as temperature sensitivity or formamide sensitivity such as temperature sensitivity or formamide sensitivity such as temperature sensitivity or formamide sensitivity such as temperature sensitivity or formamide sensitivity such as temperature sensitivity or formamide sensitivity such as temperature sensitivity or formamide sensitivity such as temperature sensitivity or formamide sensitivity such as temperature sensitivity or formamide sensitivity such as temperature sensitivity or formamide sensitivity such as temperature sensitivity or formamide sensitivity such as temperature sensitivity or formamide sensitivity such as temperature sensitivity or formamide sensitivity such as temperature sensitivity or formamide sensitivity such as temperature sensitivity or formamide sensitivity such as temperature sensitivity or formamide sensitivity such as temperature sensitivity or formamide sensitivity such as temperature sensitivity or formamide sensitivity such as temperature sensitivity or formamide sensitivity such as temperature sensitivity or formamide sensitivity such as temperature sensitivity or formamide sensitivity such as temperature sensitivity or formamide sensitivity such as temperature sensitivity or formamide sensitivity such as temperature sensitivity or formamide sensitivity such as temperature sensitivity or formamide sensitivity such as temperature sensitivity or formamide sensitivity such as temperature sensitivity or formamide sensitivity such as temperature sensitivity or formamide sensitivity such as temperature sensitivity or formamide sensitivity such as temperature sensitivity or formamide sensitivity such as temperature sensitivity or formamide sensitivity such as temperature sensitivity or formamide sensitivity such as temperature sensitivity or formamide sensitivity such as temperature sensitivity or formamide sensitivity such as temperature sensitivity or formamide sensitivity such as temperature sensitivity or formamide sensitivity.

Discussion

Herein, we confirm the recent discovery of association between pathogenic germline variation in POLR2A and a phenotypically heterogeneous neurodevelopmental disorder.6 We report the transmission of a potentially pathogenic POLR2A variant within a family—individual 8, inheriting a p.Gln1814Valfs99ter variant from a mother with a remarkably mild presentation of delayed speech and mild learning difficulties. We observe several previously unreported phenotypes (as compared to Hajies et al.9) in individuals with POLR2A-related disorders, including ataxia (observed in 7/12, or 58.3% of individuals), joint hypermobility (6/12, 50%), short stature (5/12, 41.7%), skin abnormalities including easy scarring and keratosis pilaris (5/12, 41.7%), recurrent febrile illness of unknown etiology (4/12, 33.3%), congenital cardiac abnormalities (3/12, 25%), immune system abnormalities (3/12, 25%), hip dysplasia (2/12, 16.7%), and short Achilles tendons (2/12, 16.7%). We also report a significantly higher proportion of individuals with epilepsy (8/12, 66.7%) than previously reported (3/15, 20%) (p value = 0.014196; chi-square test) and a somewhat lower proportion of individuals with hypotonia (8/12, 66.7%) than previously reported (14/15, 93.3%) (p value = 0.076309). We describe the facial dysmorphology of a subset of affected individuals, which is generally mild and nonspecific across individuals with different variants but remarkably similar for the two reported individuals sharing the same variant (individuals 4-5) (Figure 2).

In this cohort, previously unreported neuroradiological anomalies include polymicrogyria (2/12, 16.7%) and various benign, congenital anomalies, which cannot yet be ruled out as unrelated to POLR2A dysfunction, each occurring in a single individual: Rathke cleft cyst, hemangioma, and a small, enhancing developmental venous anomaly (DVA). We also report one individual (individual 1) with a germline variant identical to a previously reported meningioma-causing somatic mutation (p.His439_Leu440del).5 As individual 1 died during infancy, the extent of correlation between germline inheritance of p.His439_Leu440del (or other pathogenic germline variants) and risk of developing meningioma remains unclear.

Due to the centrality of POLR2A in transcriptional networks and the wide range of ways in which its function is known to be regulated, it can be reasonably inferred that a spectrum of possible pathogenic genetic variants will present with differential phenotypic presentation and severity. Future efforts should focus on elucidating the molecular mechanisms of pathogenicity—common or distinct—across the spectrum of known pathogenic POLR2A variants. Phenotypes of tested mutants in yeast in most instances were relatively weak, though a subset are strongly predicted to have protein structural or stability defects. Known catalytic mutants identified in yeast cause widespread transcriptional defects when introduced into human cells and likely would not be viable in an organism.26–28 For example, a known slow-elongating variant has been introduced into mouse embryonic stem cells (Polr2a Arg749His, analogous to Arg726His in yeast), could not be transmitted through the germline, and caused early embryonic lethality.29 Mouse embryonic stem cells containing Arg749His showed defects upon neuronal differentiation, likely deriving from observed altered elongation rate, gene expression, and alternative splicing changes. Of interest, long genes, which are enriched among neuronally expressed genes, might be predicted to be especially sensitive to altered Pol II elongation or cotranscriptional splicing defects.30 Hajies et al.5 delineate two potential molecular mechanisms of disease:
haploinsufficiency (with a relatively mild presentation) and a dominant-negative effect caused by aberrant Pol II elongation. We emphasize that despite the relatively mild presentation of all individuals harboring truncating and frameshift variants reported to date (p.Gln700*, p.Gln735*, p.Pro1767fs, and p.Gln1814fs), the potentially pathogenic CTD frameshift variants (p.Pro1767fs and p.Gln1814fs) most likely escape nonsense-mediated decay and thus exhibit a dominant-negative mechanism of pathogenicity. Therefore, while haploinsufficiency associated with p.Gln700* and p.Gln735* almost certainly constitutes a distinct mechanism of pathogenicity,
expressed mutant POLR2A products can cause the full range of phenotypic severity observed in POLR2A-related disorders. Taken together, these molecular and phenotypic data suggest that these pathogenic variants constitute a spectrum of transcriptional dysfunction, with phenotypes likely explained by a combination of specific POLR2A variation in conjunction with the genetic burden across a potential variant- or domain-specific network of interacting partners. Such a model could explain both a degree of phenotypic convergence (i.e., similar facial dysmorphism in individuals 4 and 5) and variable expressivity (i.e., microcephaly, cardiac and immune abnormalities present in individual 4 but not in individual 5) or incomplete penetrance (i.e., the mother of individual 8 exhibiting only sub-clinical phenotypes) for a given disease-associating variant. To assess this model, global transcriptional profiling (via RNA-sequencing [RNA-seq], GRO-seq, etc.) could be evaluated across biologically relevant cell types or tissues using patient-derived induced pluripotent stem cell (iPSC) lines. The function of individual pathogenic variants—in the appropriate genetic background—could be assessed by transcriptional profile comparison of patient-derived cells against a split of the same cell line with a mutationally induced wild-type POLR2A. The function of different pathogenic variants could then be compared by normalizing their impact against their respective isogenic controls.

Data and Code Availability
POLR2A variants are available on ClinVar: SUB7404119, accession pending review.

Supplemental Information
Supplemental Information can be found online at https://doi.org/10.1016/j.xhgg.2020.100014.

Acknowledgments
This work was supported in part by grants UM1 HG008898 from the National Human Genome Research Institute (NHGRI) to the Baylor College of Medicine for Common Disease Genetics and UM1 HG006542 from the National Heart, Lung, and Blood Institute (NHLBI) and NHGRI to the Baylor Hopkins Center for Mendelian Genomics. C.D.K. was supported by grants R01 GM097260 and R01 GM120450 National Institute of General Medical Sciences (NIGMS). A.W.H. was supported in part by NIH T32 GM08307-26, The Cullen Foundation, and the Baylor College of Medicine’s President’s Circle. D.P. was supported by Clinical Research Training Scholarship in Neuromuscular Disease partners by the American Academy of Neurology (AAN), American Brain Foundation (ABF) and Muscle Study Group (MSG), and the International Rett Syndrome Foundation (IRSF grant #3701-1). S.M.W. is supported by the Victorian Government’s Operational Infrastructure Support Program. S.D. was supported by the National Human Genome Research Institute grant R01 HG009141. We thank Ali Jalali and Amy L. McGuire for the valuable discussions. Last, we thank the individuals participating in research and their families for their assistance and significant contributions to this research.

Declaration of interests
J.R.L. has stock ownership in 23andMe, is a paid consultant for Regeneron Pharmaceuticals, and is a co-inventor on multiple US and European patents related to molecular diagnostics for inherited neuropathies, eye diseases, and bacterial genomic fingerprinting. The Department of Molecular and Human Genetics at Baylor College of Medicine derives revenue from clinical genetic testing offered in the Baylor Genetics Laboratory.

Received: September 5, 2020
Accepted: November 6, 2020

Web resources
Baylor Genetics Laboratory, https://baylorgenetics.com
ClinVar, https://www.ncbi.nlm.nih.gov/clinvar/
Face2Gene tool, https://www.face2gene.com
OMIM, https://www.omim.org
The Human Phenotype Ontology, https://hpo.jax.org/app/

References
1. Jeronimo, C., Collin, P., and Robert, F. (2016). The RNA Polymerase II CTD: The Increasing Complexity of a Low-Complexity Protein Domain. J. Mol. Biol. 428, 2607–2622.
2. Harlen, K.M., and Churchman, L.S. (2017). The code and beyond: transcription regulation by the RNA polymerase II carboxy-terminal domain. Nat. Rev. Mol. Cell Biol. 18, 263–273.
3. Cramer, P., Bushnell, D.A., and Kornberg, R.D. (2001). Structural basis of transcription: RNA polymerase II at 2.8 Ångstrom resolution. Science 292, 1863-1876.
4. Cramer, P. (2019). Organization and regulation of gene transcription. Nature 573, 45–54.
5. Clark, V.E., Harmanec, A.S., Bai, H., Youngblood, M.W., Lee, T.I., Baranoski, J.F., Ercan-Sencicek, A.G., Abraham, B.J., Weintraub, A.S., Hnisz, D., et al. (2016). Recurrent somatic mutations in POLR2A define a distinct subset of meningiomas. Nat. Genet. 48, 1253–1259.
6. Hajies, H.A., Koster, M.J.E., Rehmansh, H., Li, D., Hakonarson, H., Cappuccio, G., Hancarova, M., Lehalle, D., Reardon, W., Schaefer, G.B., et al. (2019). De Novo Heterozygous POLR2A Variants Cause a Neurodevelopmental Syndrome with Profound Infantile-Onset Hypotonia. Am. J. Hum. Genet. 105, 283–301.
7. Hansen, A.W., Murugan, M., Li, H., Khayat, M.M., Wang, L., Rosenfeld, J., Andrews, B.K., Jiangiani, S.N., Coban Akdemir, Z.H., Sledzeck, F.J., et al.; Task Force for Neonatal Genomics (2019). A Genocentric Approach to Discovery of Mendelian Disorders. Am. J. Hum. Genet. 105, 974–986.
8. Traynelis, J., Silk, M., Wang, Q., Berkovic, S.F., Liu, L., Ascher, D.B., Balding, D.J., and Petrovski, S. (2017). Optimizing genomic medicine in epilepsy through a gene-customized
19. Chapman, R.D., Conrad, M., and Eick, D. (2005). Role of the mammalian RNA polymerase II C-terminal domain (CTD) nonconsensus repeats in CTD stability and cell proliferation. Mol. Cell. Biol. 25, 7665–7674.
20. Gibbs, E.B., Lu, F., Portz, R., Fisher, M.J., Medellin, B.P., Laremore, T.N., Zhang, Y.J., Gilmour, D.S., and Showalter, S.A. (2017). Phosphorylation induces sequence-specific conformational switches in the RNA polymerase II C-terminal domain. Nat. Commun. 8, 15233.
21. Kaplan, C.D., Jin, H., Zhang, I.L., and Belyanin, A. (2012). Dissection of Pol II trigger loop function and Pol II activity-dependent control of start site selection in vivo. PLoS Genet. 8, e1002627.
22. Malik, I., Qiu, C., Snively, T., and Kaplan, C.D. (2017). Wide-ranging and unexpected consequences of altered Pol II catalytic activity in vivo. Nucleic Acids Res. 45, 4431–4451.
23. Berroteran, R.W., Ware, D.E., and Hampsey, M. (1994). The sua8 suppressors of Saccharomyces cerevisiae encode replacements of conserved residues within the largest subunit of RNA polymerase II and affect transcription start site selection similarly to sua7 (TFIIB) mutations. Mol. Cell. Biol. 14, 226–237.
24. Kaplan, C.D. (2013). Basic mechanisms of RNA polymerase II activity and alteration of gene expression in Saccharomyces cerevisiae. Biochim. Biophys. Acta 1829, 39–54.
25. Qiu, C., Jin, H., Vedenskaya, I., Llenas, J.A., Zhao, T., Malik, I., Visibisky, A.M., Schwartz, S.L., Cui, P., Cabart, P., et al. (2020). Universal promoter scanning by Pol II during transcription initiation in Saccharomyces cerevisiae. Genome Biol. 21, 132.
26. Fong, N., Saldi, T., Sheridan, R.M., Cortazar, M.A., and Bentley, D.L. (2017). RNA Pol II Dynamics Modulate Co-transcriptional Chromatin Modification, CTD Phosphorylation, and Transcriptional Direction. Mol. Cell. Biol.
27. Fong, N., Kim, H., Zhou, Y., Ji, X., Qiu, J., Saldi, T., Diener, K., Jones, K., Fu, X.D., and Bentley, D.L. (2014). Pre-mRNA splicing is facilitated by an optimal RNA polymerase II elongation rate. Genes Dev. 28, 2663–2676.
28. Saldi, T., Fong, N., and Bentley, D.L. (2018). Transcription elongation rate affects nascent histone pre-mRNA folding and 3' end processing. Genes Dev. 32, 297–308.
29. Maslon, M.M., Braunschweig, U., Aitken, S., Mann, A.R., Kilanowski, F., Hunter, C.J., Blencowe, B.J., Kornblihtt, A.R., Adams, I.R., and Cáceres, J.F. (2019). A slow transcription rate causes embryonic lethality and perturbs kinetic coupling of neuronal genes. EMBO J. 38, e101244.
30. Zylka, M.J., Simon, J.M., and Philpot, B.D. (2015). Gene length matters in neurons. Neuron 86, 353–355.