Microduplication of the *ARID1A* gene causes intellectual disability with recognizable syndromic features

Marie Bidart, PharmD, PhD[1,2], Michèle El Atifi, PhD[1,2], Sarra Miladi, MS[1,2],
John Rendu, PharmD, PhD[2,3], Véronique Satre, MD, PhD[2,4,5], Pierre F. Ray, PhD[2,3,5],
Caroline Bosson, PharmD[3], Françoise Devillard, MD[4], Daphné Lehalle, MD[6], Valérie Malan, MD, PhD[7],
Jeanne Amiel, MD, PhD[6], Maria Antonietta Mencarelli, MD, PhD[8], Margherita Baldassarri, MD[8,9],
Alessandra Renieri, MD, PhD[8,9], Jill Clayton-Smith, MD, PhD[10], Gaëlle Vieville, MS[4],
Julien Thevenon, MD, PhD[11], Florence Amblard, MD[4], François Berger, MD, PhD[1,2],
Pierre-Simon Jouk, MD, PhD[2,4] and Charles Coutton, PharmD, PhD[2,4,5]

Purpose: To determine whether duplication of the *ARID1A* gene is responsible for a new recognizable syndrome.

Methods: We describe four patients with a 1p36.11 microduplication involving *ARID1A* as identified by array–comparative genomic hybridization . We performed comparative transcriptomic analysis of patient-derived fibroblasts using RNA sequencing and evaluated the impact of *ARID1A* duplication on the cell cycle using fluorescence-activated cell sorting. Functional relationships between differentially expressed genes were investigated with ingenuity pathway analysis (IPA).

Results: Combining the genomic data, we defined a small (122 kb), minimally critical region that overlaps the full *ARID1A* gene. The four patients shared a strikingly similar phenotype that included intellectual disability and microcephaly. Transcriptomic analysis

revealed the deregulated expression of several genes previously linked to microcephaly and developmental disorders as well as the involvement of signaling pathways relevant to microcephaly, among which the polo-like kinase (PLK) pathway was especially notable. Cell-cycle analysis of patient-derived fibroblasts showed a significant increase in the proportion of cells in G1 phase at the expense of G2-M cells.

Conclusion: Our study reports a new microduplication syndrome involving the *ARID1A* gene. This work is the first step in clarifying the pathophysiological mechanism that links changes in the gene dosage of *ARID1A* with intellectual disability and microcephaly.

Genet Med advance online publication 1 December 2016

Key Words: 1p36.11 microduplication; *ARID1A*; intellectual disability; microcephaly; RNA sequencing

INTRODUCTION

ARID1A (AT-rich interactive domain-containing protein 1A), also known as BAF250A, is one of the main subunits of the switch/sucrose nonfermentable (SWI/SNF)-like chromatin-remodeling complex (BAF complex), a multiprotein complex essential for chromatin remodeling. This complex plays a key role in the epigenetic regulation of genome expression and has been implicated in several cellular functions, including proliferation, self-renewal of stem cells, DNA methylation, and damage repair.[1] In particular, the ARID1A-containing chromatin-remodeling complex regulates processes related to tumor suppression,[2] and somatic mutations that inactivate

the *ARID1A* gene have been identified in a wide spectrum of human neoplasms.[3] Moreover, *ARID1A* shows one of the highest mutation rates across different types of human cancer.[4]

Recent discoveries of *ARID1A* germ-line mutations associated with Coffin-Siris syndrome (CSS; OMIM 135900) demonstrated that ARID1A also plays a critical role in human neuronal development.[5] CSS is a rare autosomal-dominant syndrome characterized by intellectual disability, growth deficiency, microcephaly, coarse facial features, and hypoplastic nails, especially on the fifth finger and/or toe. Other findings commonly include feeding difficulties, ophthalmologic abnormalities, brain malformations, and hearing loss.[6,7] Large-scale

studies reported that 41 to 100% of CSS subjects carry a germline mutation in one of the BAF complex subunit genes.[5,8–12] Overall, mutations in the *ARID1A* gene are found in only 7.5% of cases, a lower frequency than that observed for the other genes of the BAF complex.[13] All CSS-causing mutations previously identified in *ARID1A* are heterozygous and have a loss-of-function effect, indicating that ARID1A is a dosage-sensitive gene.[5] Therefore, it raises the possibility that the gain of a copy of *ARID1A* could also be pathogenic. We previously identified a microduplication encompassing the *ARID1A* gene in a girl presenting with intellectual disability and microcephaly.[14] In the present report, we compare the genomic data and clinical features of four patients carrying a small 1p36.11 microduplication encompassing the complete *ARID1A* gene. Subsequent molecular studies then allowed us to explore the functional pathways and biological functions disturbed by *ARID1A* duplication.

MATERIALS AND METHODS

Patients
Following the diagnosis of our first patient,[14] three additional patients were identified as harboring small 1p36.11 duplications that fully included *ARID1A*; they were recruited through the DECIPHER database. To limit the number of genes encompassed to be able to establish a precise genotype–phenotype correlation, only duplications smaller than 1 Mb were selected. Clinical information was gathered through a detailed questionnaire addressed to the referring clinician. An overview of these four patients (PT1, PT2, PT3, and PT4) and their clinical characteristics are presented in **Table 1**. A more detailed clinical report for each patient can be found in the **Supplementary Clinical Information** online. Photographs of patients (PT1–PT4) are shown in **Figure 1a**. Two patients were from France (PT1 and PT2), one from Italy (PT3), and one from the United Kingdom (PT4). Written informed consent for publication of their clinical details and/or clinical images was obtained from all patients and institutional approval was given by the local medical ethical committee (Grenoble Institutional Review Board, 6705). We also included three healthy volunteers (C1–C3) without *ARID1A* duplication (data not shown) as control samples. Control subjects were recruited in France with self-declared European ancestry. Two were females, aged 18 and 25, respectively, and one was a 22-years-old male. Written informed consent was obtained from all controls.

Chromosomal microarray analysis
DNA was extracted from peripheral blood lymphocytes of the patients following standard protocols. All four 1p36.11 microduplications were detected using a 60,000-oligonucleotide microarray (Human Genome CGH Microarray Kit 60K; Agilent Technologies, Santa Clara, CA) according to the manufacturer's instructions. The Database of Genomic Variants (http://projects.tcag.ca/ variation/) was used to compare our findings to those of previously reported studies. Genomic positions are based on the UCSC GRCh37/hg19 assembly. Verifications of gain of 1p36.11 material in the proband and

parental testing were subsequently performed as described in the **Supplementary Methods** online.

Skin biopsies
Skin biopsies were performed for PT1, PT2, and PT3 and the three control individuals (C1–C3). Samples were obtained from a cutaneous biopsy of the forearm skin using a scalpel and biopsy forceps (size ≈0.5 mm³). Only fibroblast cultures from PT1 and PT2 were included in our study because fibroblasts from PT3 did not multiply in vitro.

Cell culturing
Biopsy tissue was rinsed several times with an antibiotic–antimycotic solution. Tissue was then placed in 1 ml of collagenase solution and incubated for 24 h at 37 °C in a 5% CO_2 atmosphere. The resulting confluent culture was designated passage 1. Cells were then split to generate three 25-cm² flasks. Cultures at 70% confluence were harvested by centrifugation (340*g*), and cell pellets were stored at −80 °C in RNAlater (Ambion, Cambridge, MA).

Next-generation sequencing and data analysis
RNA paired-end sequencing (RNA-seq) was performed for two patients (PT1 and PT2) and three controls with RNA from skin fibroblasts. The functional relationships between genes differentially expressed between patients and controls were subsequently investigated using Ingenuity Pathway Analysis (IPA) software (Qiagen, Hilden, Germany, https://analysis.ingenuity.com). Further details are provided in the **Supplementary Methods** online.

Quantitative reverse-transcription polymerase chain reaction
Quantitative reverse-transcription polymerase chain reaction (qRT-PCR) experiments were first performed using two candidate genes: *ARID1A* and *PIGV*. Then, the expression of the various deregulated genes was verified by qRT-PCR to confirm the results observed with RNA sequencing. Validation was applied to a subset of 24 genes that were differentially expressed by at least fourfold in patient-derived fibroblasts using RNA-seq. Because the mitotic role of polo-like kinase pathway was among the most statistically significant pathways deregulated in our patients, we decided to test the expression of *PLK1*, another key gene of the PLK pathway, which was not initially detected with RNA-seq in PT1, probably owing to technical failure. qRT-PCR was performed as described in the **Supplementary Methods** online. The primer sequences and qRT-PCR protocol are presented in **Supplementary Table S1** online.

Western blot analysis
We performed western blot analyses of ARID1A and PLK4 in patient and control fibroblasts. Total protein was extracted from cells in RIPA lysis buffer (R-078; Sigma-Aldrich, St. Louis, MO). Western blotting was performed on NuPAGE Novex 4–12% Bis-Tris Mini Gels (Thermo Fisher Scientific, Waltham, MA). The membranes were incubated with primary antibody

Table 1 Comparison of the clinical features and molecular data of the four patients presented in this study

	Patient 1	Patient 2	Patient 3	Patient 4	
Sex (M/F)	F	M	F	F	
Age at last examination (years)	6	13	34	6	
Duplication coordinates	Chr1: 27001256-27190935	Chr1: 26978354-27256595	Chr126796023-27123236	Chr1: 26917016-27169244	
Deletion size (Kb)	190	278	327	252	
Encoding genes present	ARID1A, PIGV, *ZDHHC18*, and *SFN*	ARID1A, PIGV, *ZDHHC18*, *SFN*, GPN2, GPATCH3, NR0B2, and NUDC	DHDDS, HMGN2, RPS6KA1, ARID1A, and PIGV	ARID1A, PIGV, and *ZDHHC18*	
Inheritance	*De novo*	*De novo*	*De novo*	*De novo*	
Clinical and developmental findings					Total
Neurodevelopment					
Developmental delay	+	+	+	+	4/4
Language delay	+	+	+	+	4/4
Hypotonia	-	+	+	-	3/4
Microcephaly	+ (45 cm, <p3)	+ (52 cm, <p10)	+ (47 cm, <p3)	+ (49 cm, <p3)	4/4
Stereotyped behavior	+	+	+	-	3/4
Epilepsy/seizures	-	-	-	-	0/4
Ectodermal					
Nail anomalies	-	-	+	-	1/4
Sparse scalp hair	+	+	-	+	3/4
Abnormal dentition	-	-	+	+	2/4
Facial					
Eye anomalies	-	-	Deep-set eyes	Iris mesodermal defects	2/4
Broad nasal bridge/tip	+	-	+	+	3/4
Prominent/low columella	-	-	+	+	2/4
Short philtrum	+	-	+	-	2/4
Thin upper lip	+	-	+	+	3/4
Small mouth	-	-	+	+	2/4
Ear anomalies	+	+	-	-	2/4
Skeletal					
Growth restriction	+	+	+	-	3/4
Hand and foot anomalies	Postaxial hexadactyly, overlapping toes	Short, clinodactyly of 5th finger, flat valgus feet	Short, clinodactyly of 5th finger, abnormally inserted thumbs, flat valgus feet	Flat feet, clinodactyly of 5th finger	4/4
Spinal malformations	-	Dorsal kyphosis	Cervical-dorsal-lumbar scoliosis, dorsal kyphosis, lumbar lordosis	-	2/4
Gastrointestinal					
Feeding problems	+	-	+	+	3/4
GERD	+	-	+	-	2/4
Constipation	+	+	+	+	4/4
Other clinical features					
	Cold hand and feet syndrome	Recurrent airway infections, PA-VSD, short neck, vasomotor problems	Recurrent airway infections, edema of lower limbs, centripetal obesity	Recurrent airways infections, hoarse voice	

Coordinates of all variations are based on the UCSC GRCh37/hg19 assembly.

+, Feature present; -, feature absent; GERD, gastroesophageal reflux disease; p, percentile; PA-VSD, pulmonary atresia with ventricular septal defect.

(anti-ARID1A: ab71394; Abcam, Cambridge, UK; anti-PLK4: ab71394; Abcam).

siRNA-mediated ARID1A silencing in fibroblast cells
At 1 day after plating, fibroblasts derived from PT1 and PT2 were transfected with 10 nmol/l of ARID1A-targeted short interfering RNA (siRNA), using Lipofectamine RNAiMAX Reagent (Life Technologies, Carlsbad, CA) according to the manufacturer's recommendations. ARID1A-targeted siRNA was purchased from Ambion (AM16708, ID 147403; Life Technologies). A Blast search on the NCBI database ensured specific targeting to ARID1A mRNA. We used Silencer Select
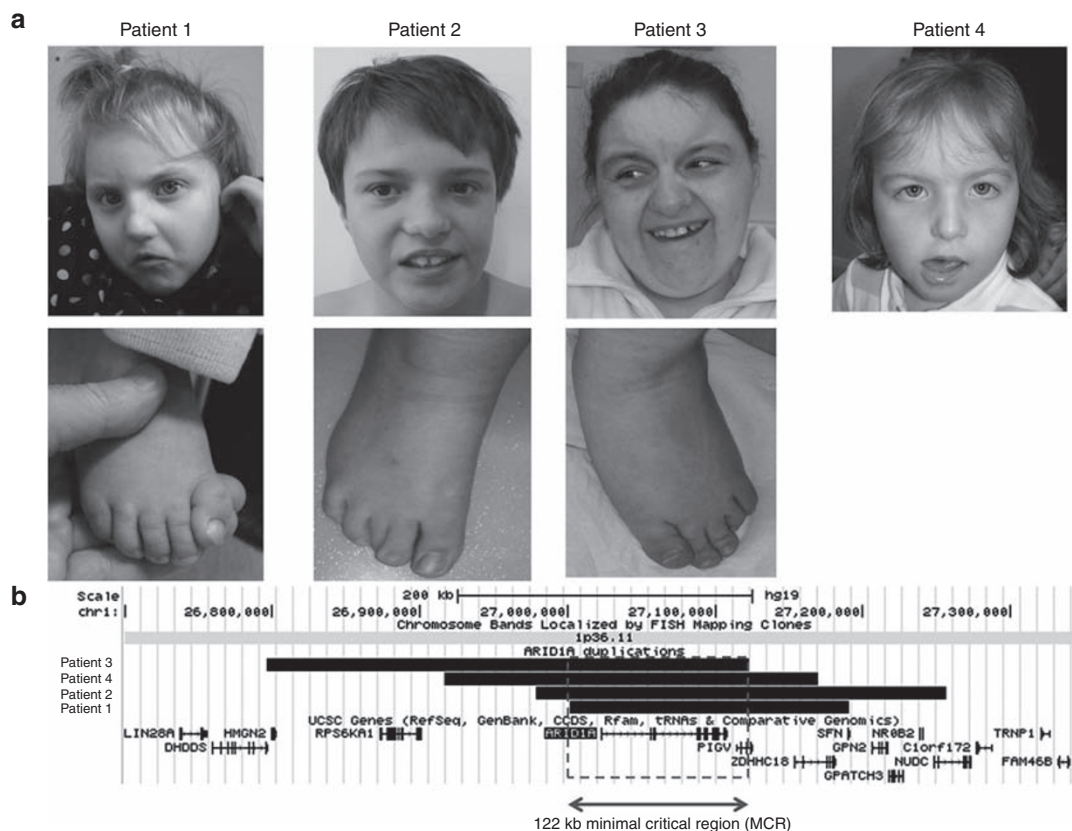
Figure 1 Photographs and genomic rearrangements in patients with *ARID1A* duplications (PT1–PT4). (**a**) Patient 1 (PT1), 6 years old; patient 2 (PT2), 13 years old; patient 3 (PT3), 34 years old; and patient 4 (PT4), 10 years old. All patients presented with intellectual disability and microcephaly and have a characteristic nasal tip with notched alae nasi and a low columella. The feet are very distinctive, with brachydactyly and shortening and medial deviation of the great toes. (**b**) Alignment of the genomic rearrangements reported in the four patients (PT1–PT4) with 1p36.11 duplications. A small (122 kb), minimally critical region (MCR; gray dash line box) (chr1:27,001,256-27,123,236) common to all four index cases was defined. Among the duplicated genes in the MCR, only the *ARID1A* gene was completely duplicated. A second gene, *PIGV*, was partially included.

Negative Control No. 1 siRNA (AM17140; Life Technologies) as scrambled nontargeting siRNA (negative control). RNA extractions were performed 48 h after siRNA transfection. Finally, we performed subsequent qRT-PCR for a subset of 10 PLK pathway genes that were differentially expressed by at least fourfold in patient-derived fibroblasts using RNA-seq.

Cell-cycle analysis

We performed cell-cycle analysis for patients PT1 and PT2 and for the control pool (C1–C3). A total of 5×10^6 cells were harvested, rinsed with cold phosphate-buffered saline (PBS), and fixed with absolute ice-cold ethanol at −20 °C. Fixed cells were rinsed with cold PBS, followed by incubation with PBS containing 500 μg/ml propidium iodide and 1 mg/ml RNase A for 1 h at 37 °C. The DNA content of labeled cells was acquired using fluorescence-activated cell sorting with a BD FACSCalibur flow cytometer (BD Biosciences, Le Pont de Claix, France). Each experiment was performed in triplicate.

RESULTS

CMA analysis

Array–comparative genomic hybridization analyses of the four patients showed a genomic duplication in chromosome band

1p36.11. For all these patients, no other chromosomal anomalies larger than three probes were observed, excluding well-known benign copy-number variations. All duplications were confirmed and occurred de novo. These aberrations ranged in size from 190 to 327 kb and contained three to eight genes (**Figure 1b**). All aberrations led to a complete duplication of *ARID1A*. Together, these duplications defined a minimal critical region (MCR) of 122 kb (chr1:27,001,256-27,123,236) that fully included the *ARID1A* gene and partially included the *PIGV* gene (**Figure 1b**).

Overview of differentially expressed genes induced by *ARID1A* microduplication

The overview of differentially expressed genes is based on the analysis of two patients, PT1 and PT2, and three controls. We identified 2,890 differentially expressed genes between patients and controls, including 1,325 downregulated genes (fold change <0.5) and 1,535 upregulated genes (fold change >2). According to our IPA criteria, we identified five top canonical pathways in which the uploaded gene set participated more than could be expected by chance ($P > 3$): cell-cycle control of chromosomal replication ($P = 7.37$), hereditary breast cancer signaling

(P = 5.48), role of BRCA1 in DNA damage response (P = 5.29), mitotic role of polo-like kinase (P = 4.96), and ataxia telangiectasia mutated protein signaling (P = 3.54). Of these top five pathways, only the mitotic role of the polo-like kinase pathway showed a significant absolute z-score >2 (2.673) (**Figure 2a**).

Next, we examined diseases and biological functions overrepresented in our data. We classified items according to the best P-value with a significant absolute z-score ($\geq$2) (**Supplementary Table S2** online). The top-ranked biological functions included the general categories of cell death and
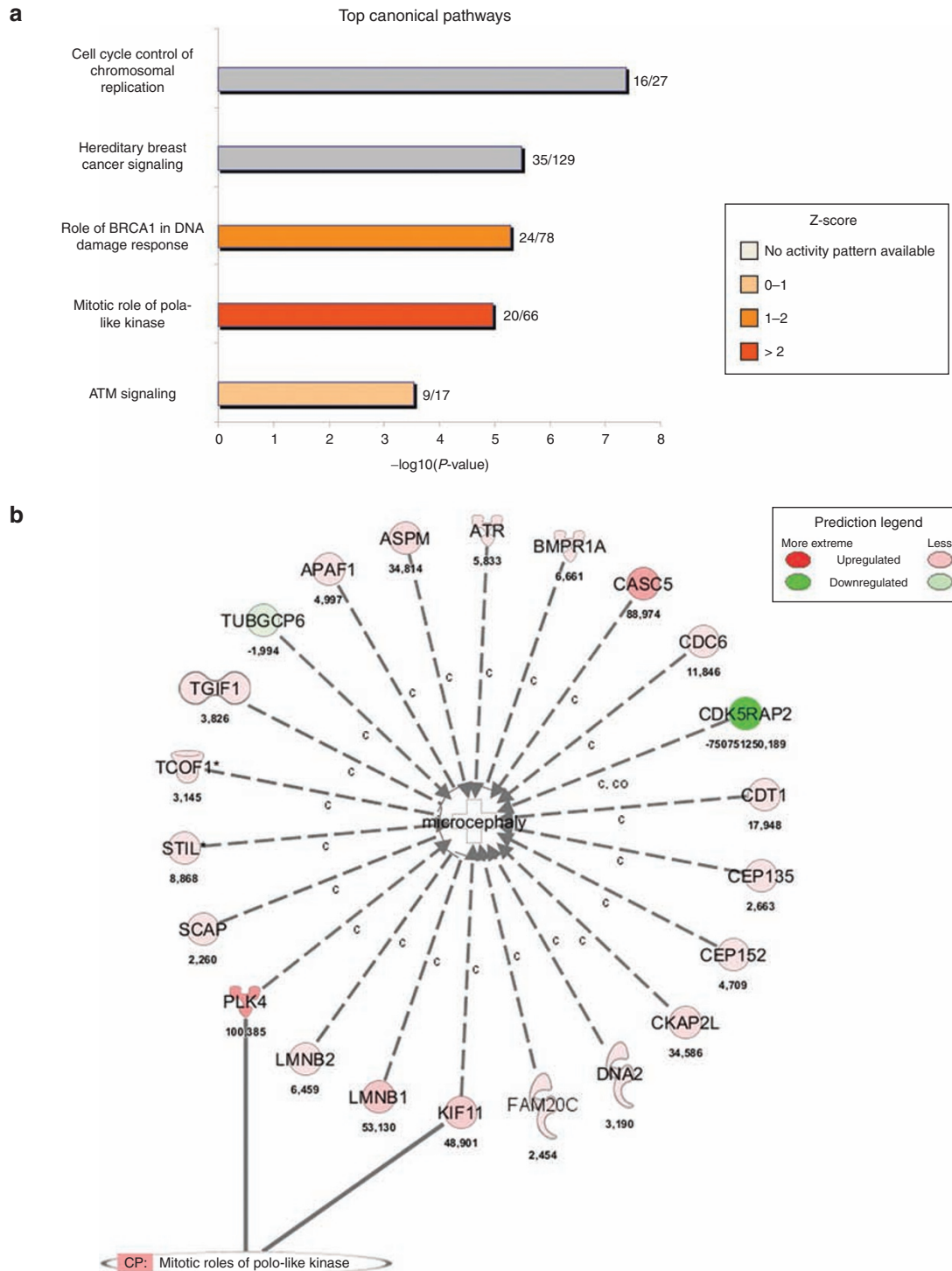

Figure 2 RNA sequencing results. (**a**) The five canonical pathways identified by ingenuity pathway analysis (IPA) as significantly (P >3) associated with the set of genes deregulated between patients and controls. The absolute z-score as well as the number of genes differentially expressed among the total genes assigned to the canonical pathway by IPA are indicated. (**b**) Variations in the expression of genes implicated in microcephaly observed in patient-derived fibroblasts. Red genes are upregulated genes and green genes are downregulated genes.

survival, cell cycle, and cell growth and proliferation, as well the more specific category of DNA replication, recombination, and repair. Detailed biological functions are presented in **Supplementary Table S2** online. Regarding diseases, the top-ranked categories were cancer, which showed high levels of enrichment (up to $P = 2.51e^{-10}$), and developmental disorders. Of these developmental disorders, IPA predicted a significant enrichment of genes related to congenital malformations of the skeleton ($P = 4.13e^{-6}$, z-score: 2.146), craniofacial abnormalities ($P = 5.49e^{-6}$, z-score: 2.176), growth failure ($P = 6.16e^{-5}$, z-score: 4.002), microcephaly ($P = 8.54e^{-5}$, z-score: 2.630), and congenital malformations of the brain ($P = 2.76e^{-4}$, z-score: 2.534) (**Supplementary Table S2** online). We subsequently investigated the microcephaly gene network and observed that the polo-like kinase 4 (*PLK4*) gene was the most strongly deregulated gene (**Figure 2b**). We also noted that the *PLK4* gene

harbors the strongest overexpression in the mitotic role of the polo-like kinase pathway (**Supplementary Table S3** online).

Validation of the differential expression of selected genes by qRT-PCR

qRT-PCR was performed for the two genes (*ARID1A* and *PIGV*) in the 122-kb MCR. Only the expression of *ARID1A* increased approximately 1.5-fold compared to controls (**Supplementary Table S4** online), whereas *PIGV* expression did not seem to be altered. qRT-PCR validation was also applied to a subset of 25 genes, including *PLK1*. Eleven of the 25 selected genes belonged to the mitotic role of polo-like kinase pathway, and each showed significantly increased expression in patient-derived fibroblasts compared to control-derived fibroblasts (**Figure 3a** and **Supplementary Table S4** online). We confirmed that *PLK4* expression appeared strongly upregulated in the two patients
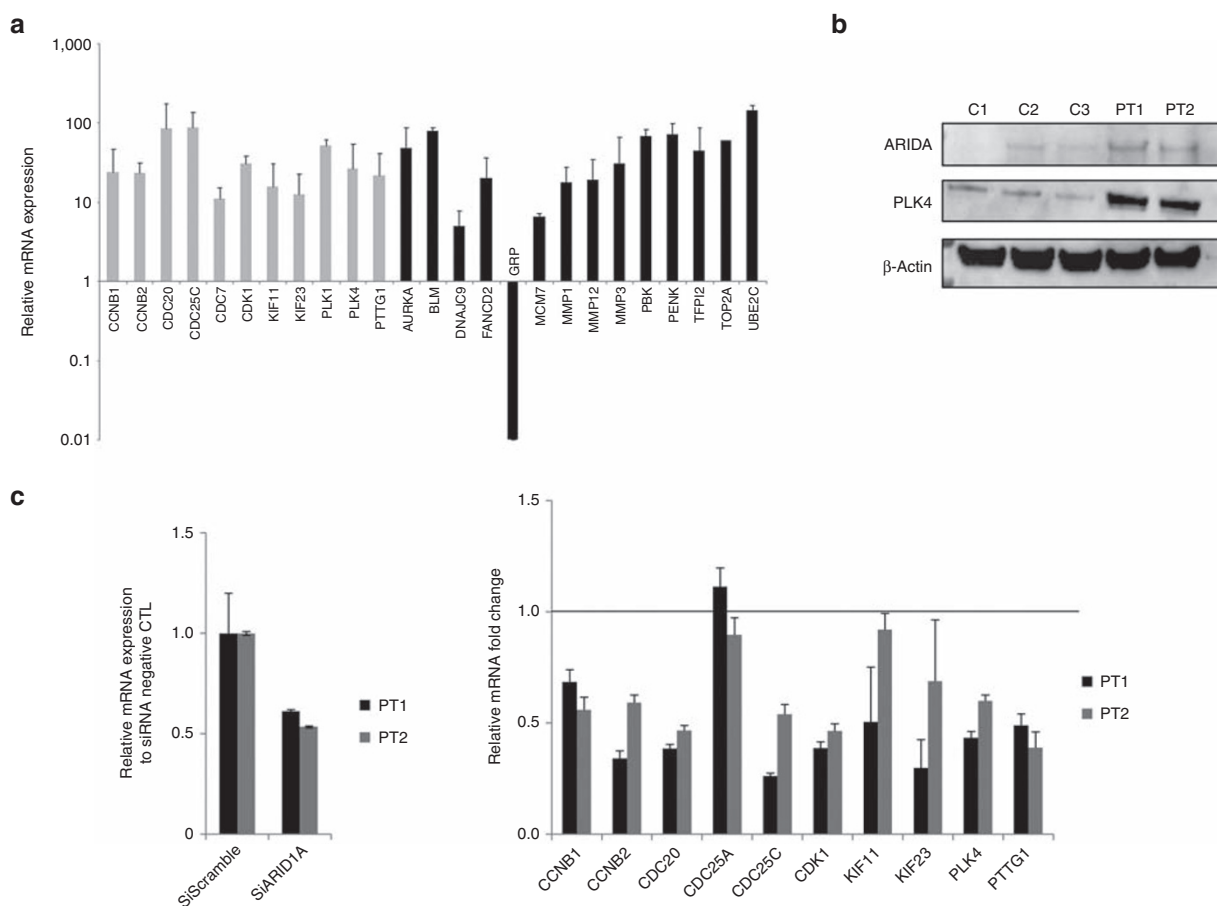

Figure 3 Validation of deregulated genes derived from RNA-seq analysis. (**a**) Validation of gene expression values derived from RNA-seq by qRT-PCR. qRT-PCR was applied to a subset of 25 genes. Eleven of the 25 selected genes belonged to the mitotic role of polo-like kinase pathway (gray bars). ARN samples from patient 1 (PT1) and patient 2 (PT2) were tested independently and compared to the control pool (C1–C3). Data are presented as the mean of the experiments ± standard deviation for the two patients. qRT-PCR data were normalized using two reference genes, ribosomal protein L27 (*RPL27*) and glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*), with the -ΔΔCt method. (**b**) Western blot analyses of ARID1A and PLK4 in patient and control fibroblasts. Endogenous expression of ARID1A and PLK4 were examined in PT1 and PT2 and control fibroblasts (C1–C3) by western blot. Western blot analyses showed that ARID1A expression is slightly increased, whereas PLK4 is markedly increased in PT1 and PT2 cells. Beta-actin served as a loading control. (**c**) siRNA experiments targeting *ARID1A* in patient-derived fibroblasts. The left panel shows the relative gene expression of *ARID1A* after siRNA transfection in PT1 and PT2 fibroblasts. The right panel shows the relative gene expression of the subset of PLK pathway genes after siRNA targeted against the *ARID1A*. The gray line indicates the normalized level of negative siRNA control. Expression of *GAPDH* and *RPL27* was used as a normalization control. Data were pooled from two individual experiments containing three replicates per condition. Error bars show SEM.

with *ARID1A* duplication (mean FC: 31) (**Figure 3a** and **Supplementary Table S4** online). We also tested the expression of PLK1, which was initially not detected with RNA-seq in PT1. qRT-PCR showed a strong increase in *PLK1* expression in PT1 (FC: 44) and in PT2 (FC: 59). qRT-PCR experiments were performed for 14 other genes assigned to different pathways and confirmed the significant gene expression alterations observed with RNA sequencing. We were able to identify a disparity in the fold-change values calculated using qRT-PCR and RNA-seq. This could be due to the normalization process of data analysis. RNA-seq allows absolute quantitation of mRNA levels, whereas our quantitative PCR experiments used relative quantification. qRT-PCR is able to detect the main mRNA and some specific splice variants of our selected genes, but additional variants should exist; they can be detected and quantified only by RNA-seq. Despite such differences in the fold-change values, our data showed that the qRT-PCR results were consistent with the RNA-seq results for all 25 selected genes, confirming that comparative RNA-seq analysis is a reliable quantitative approach.

Western blot analyses
Western blot analyses showed that ARID1A and PLK4 protein expressions were increased cells in PT1 and PT2 (**Figure 3b**). ARID1A protein expression was slightly but significantly increased in patient cells, which is consistent with the 1.5-fold overexpression detected by quantitative PCR. PLK4 expression was markedly increased in patient cells, confirming the strong *PLK4* overexpression evidenced by the quantitative PCR and RNA-seq data.

In vitro inhibition of *ARID1A* expression by siRNA
ARID1A-targeted siRNA experiments were performed in cultured cells from PT1 and PT2 and led to a twofold decrease of *ARID1A* expression (**Figure 3c**). The decrease in the expression of *ARID1A* caused significant downregulation of the expression of almost all PLK pathway genes tested by

qRT-PCR (**Figure 3c**). This result demonstrated that variations of *ARID1A* gene expression could modulate the expression of several PLK pathway genes and strongly supports the idea that genomic *ARID1A* duplication could be directly responsible for the dysregulated mitotic role of the polo-like kinase pathway observed with RNA-seq in patient fibroblasts.

Cell-cycle perturbation
We examined the cell-cycle distribution by flow cytometry on control and patient fibroblasts. The distribution of the cell-cycle phases based on analysis using BD Accuri C6 Flow Cytometry software is shown in **Figure 4**. PT1 and PT2 fibroblasts exhibited a significant increase in the fraction of G1 phase cells ($P < 0.05$) and a corresponding reduction in the proportion of G2-M phase cells ($P < 0.05$). These data suggest that *ARID1A* duplication in patient-derived fibroblasts may slow cell-cycle progression from G1 to S phase.

DISCUSSION
Somatic and germ-line mutations in SWI/SNF components have previously been implicated in tumor development.[15] Recently, heterozygous germ-line mutations in components of the BAF complex were identified in patients with nonsyndromic intellectual disability, CSS, and Nicolaides-Baraitser syndrome (OMIM 601358).[8] The common denominators of the phenotypes of these patients are severe intellectual disability and speech delay. To date, only truncating mutations leading to a loss of function of the *ARID1A* gene have been reported.[8] Here, we present four patients with 1p36.11 duplications detected by array–comparative genomic hybridization. These patients shared a similar phenotype, consisting primarily of microcephaly, intellectual disability, delayed motor milestones, hand and foot anomalies, growth impairment, constipation, frequent airway infections, dysmorphic facial features, and stereotypies. They all have a characteristic nasal tip with notched alae nasi and a low columella (**Figure 1a**). Deficiency of the lateral vermillion border of the upper lip is also seen (**Figure 1a**), as well
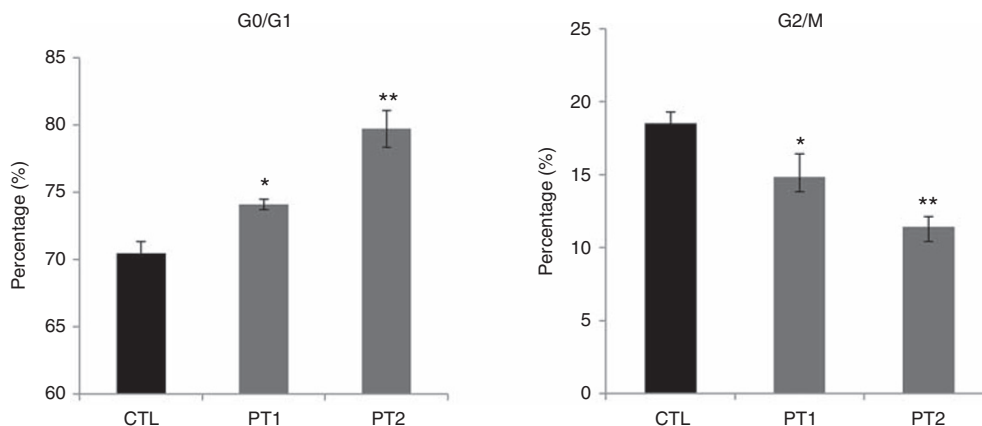

Figure 4 *ARID1A* overexpression induced G0/G1 cell-cycle arrest. Control (CTL), patient 1 (PT1), and patient 2 (PT2) cells were stained with propidium iodide and analyzed by flow cytometry. The percentage of cells in G0/G1 phase was significantly higher in the patient (gray) group than in the control group (black). A reciprocal reduction of cells in G2/M phase was observed in patient samples. Each experiment was performed in triplicate. *Significant at $P < 0.05$, *t*-test; **significant at $P < 0.01$, *t*-test.

as a high forehead and sparse hair (**Figure 1a**). The feet are very distinctive, with brachydactyly and shortening and medial deviation of the great toes (**Figure 1a**). We found a 122-kb MCR between these patients. Two genes were included in the MCR, but only the *ARID1A* gene was fully duplicated; therefore, it seemed to be the best candidate to explain the clinical features. The *PIGV* gene is only partially included, and mutations in each *PIGV* allele (homozygous or compound heterozygous) are necessary to substantially reduce the level of *PIGV* protein expression or to be deleterious.[16] In addition, *PIGV* expression did not appear to be deregulated in our patients according to RNA-seq and qRT-PCR experiments (**Supplementary Table S4** online). Therefore, partial or complete duplication of *PIGV* is unlikely to be pathogenic.[14]

Conversely, *ARID1A*—a chromatin-remodeling gene with a dose-dependent effect because haploinsufficiency can cause CSS,[5]—shows the highest likelihood of being a triplosensitive gene. Our previous and present studies revealed that *ARID1A* duplication leads to a 1.5-fold overexpression in patients (**Supplementary Table S4** online).[14] Subsequent western blot analysis was consistent with this modest gene overexpression detected by quantitative PCR and confirmed that ARID1A protein expression was slightly but significantly increased in fibroblasts from PT1 and PT2 (**Figure 3b**). Consequently, it is possible that duplication of the *ARID1A* gene could lead to a gain of function of the protein with enhanced activity. This hypothesis is supported by our present RNA-seq analyses that showed the strong deregulation of many genes involved in several biological functions (cell cycle, cell proliferation, and cell death as well DNA replication, recombination, and repair) for which ARID1A was previously demonstrated to be a key transcriptional regulator.[17]

Two larger duplications that fully include *ARID1A* are reported in the DECIPHER database. The first is an 18.9-Mb duplication (ID 306661) that encompasses 314 genes; the second (ID 30771) is a 1.8-Mb duplication that encompasses 34 genes. Clinical information was available only for patient 30771, who presented with microcephaly, motor delay, fifth-finger clinodactyly, and 2/3 toe syndactyly. Even if the genotype–phenotype correlation for such large duplications is uncertain, we can note that some clinical features presented by patient 30771, such as microcephaly and hand/foot anomalies, are common to those found in our four patients.

Finally, we note that no duplications that fully included *ARID1A* were reported in databases reporting copy-number variations from healthy subjects, such as the Database of Genomic Variants. In light of these considerations, there is strong evidence that *ARID1A* duplications are pathogenic and likely causative for the phenotype reported in the four patients.

An increasing number of observations in the literature demonstrate a close link between the constitutional alteration of tumor suppressor genes and neurodevelopmental disorders.[18,19] Nevertheless, the molecular mechanisms linking the increase of *ARID1A* gene product with the reported clinical features remain to be understood. In our study, we observed

that *ARID1A* duplication in patient-derived fibroblasts slowed cell-cycle progression from G1 to S phase (**Figure 4**). This result was consistent with the expected role of ARID1A in the cell-cycle checkpoint machinery[20] and with the body of oncological literature showing that the overexpression of *ARID1A* leads to an arrest of cell-cycle progression in G1/S phase.[21]

Interestingly, it has been demonstrated that the BAF complex is involved in the proliferation and differentiation of neural stem and progenitor cells.[8,22] Moreover, it is understood that the brain's size at birth is dependent on the ability of neuroprogenitor cells to proliferate and self-renew.[23,24] Therefore, a slight perturbation in the number of cell divisions of progenitor and stem cells, such as slowed cell-cycle progression, can have dramatic effects on brain size and may lead to microcephaly.[25] Alterations in cell-cycle control and kinetics with a prolonged G1 phase reported in patient-derived fibroblasts may thus be a major factor in intellectual disability and microcephaly. However, cell-cycle alterations shown in patient-derived fibroblasts could be different in neural progenitors and should be further evaluated.

From a molecular point of view, IPA analysis of our RNA-seq data showed significant enrichment of genes related to several developmental disorders reported in all our patients, such as congenital malformations of the skeleton ($P = 4.13\mathrm{e}^{-6}$, *z*-score: 2.146), craniofacial abnormalities ($P = 5.49\mathrm{e}^{-6}$, *z*-score: 2.176), growth failure ($P = 6.16\mathrm{e}^{-5}$, *z*-score: 4.002), microcephaly ($P = 8.54\mathrm{e}^{-5}$, *z*-score: 2.630), and congenital malformations of brain ($P = 2.76\mathrm{e}^{-4}$, *z*-score: 2.534). Most intriguingly, transcriptional analysis also showed that the majority of the known microcephaly-associated genes were upregulated in our patients' fibroblasts (**Figure 2a**, **Supplementary Table S2** online).[24] Almost all of these genes are closely linked to cell-cycle regulation and/or centrosomal functions and are involved in severe primary microcephaly.[26,27]

Primary microcephaly is defined by an occipitofrontal circumference (OFC) less than the third percentile present at birth and should be distinguished from secondary microcephaly, which develops postnatally.[28] In our patients, OFC at birth was available only for PT1, for whom it was at the 10th percentile. However, biometric parameters at the lower limit of the normal range were reported for PT2 without more details. At the time of the last clinical evaluation, three patients presented with an OFC lower than the third percentile and one with an OFC lower than the 10th percentile. Although some OFCs at birth were missing, these data may suggest that the phenotype of these patients is more likely to be secondary microcephaly. The biological and clinical impacts of the overexpression of these microcephaly-associated genes, most of which are unlikely to be dosage-sensitive, remain uncertain. However, we cannot exclude the hypothesis that overexpression of one or more of these genes could have a significant negative effect on common cellular functions or pathways but may be less dramatically harmful for the cell than the downregulation or the total absence of the encoded protein. This might partly explain why our patients present with a secondary microcephaly and not a primary microcephaly, but it remains to be demonstrated.

Beyond the overexpression of some microcephaly-associated genes, IPA revealed that the mitotic role of polo-like kinase pathway was among the most statistically significant pathways deregulated in patients with microduplications encompassing *ARID1A* (*P* = 4.96) and the only pathway with an absolute *z*-score >2 (2.673) (**Figure 2a**). Almost all altered transcripts of the PLK pathway in patients with *ARID1A* duplication were upregulated (**Supplementary Table S3** online). ARID1A-targeted siRNA experiments in patient fibroblasts confirmed the existence of a strong link between variations in *ARID1A* expression and the expression level of several PLK pathway genes (**Figure 3c**). Although we cannot totally exclude the possibility that some differences between cases and controls (e.g., in age) may induce confounding factors, these data converged and support that the genomic *ARID1A* duplication could be directly responsible for the dysregulation of the mitotic role of polo-like kinase pathway observed with RNA-seq in patient fibroblasts. In the mitotic role of polo-like kinase pathway, *PLK4* presents the highest gene overexpression detected by RNA-seq (FC >100) (**Supplementary Table S3** online). Our transcriptomic data were confirmed by qRT-PCR and western blot analysis showing that PLK4 is also markedly increased at the protein level in patient fibroblasts (**Figure 3b**).

Interestingly, upregulation of Plk4 was previously shown to impede brain development and cause microcephaly in mice.[29] Furthermore, individuals with loss-of-function mutations in *PLK4* display profound microcephaly, confirming that gene *PLK4* dosage variations (loss or gain) are pathogenic and that both lead to microcephaly.[30,31] Using qRT-PCR, we found that another PLK gene, the *PLK1* gene, was strongly overexpressed in our patients. Experiments with different animal models have demonstrated that gene dosage variations of *PLK1* can also have a critical effect on cortical neurogenesis, the regulation of mammalian brain size, and the pathogenesis of congenital neurodevelopmental disorders, such as microcephaly.[32,33] In addition, in 2013, Barber et al.[34] described two patients with intellectual disability and dysmorphic features with a 16p11.2–p12.2 duplication including *PLK1*, which seems to be one of the best candidate genes to explain the phenotype. These patients present with the most clinical features observed in our four patients with *ARID1A* duplications, such as developmental and psychomotor delay, microcephaly, growth restriction, stereotyped behaviors, dysmorphic features, ear anomalies, feet and hand anomalies, and recurrent infections.[34]

Interestingly, the reciprocal 16p11.2–p12.2 deletion leads to microcephaly, suggesting that both overexpression and decreased expression of some genes can lead to a reciprocal effect such as microcephaly or macrocephaly. This mirror effect was not observed in our case because duplications or inactivating mutations of the *ARID1A* gene lead to similar defects such as microcephaly. Moreover, many other clinical features, such as developmental delay, hand and foot anomalies, growth impairment, gastrointestinal problems, and spinal malformations, are common in CSS patients with the 1p36.11 duplication. The same observations were made for other reciprocal deletions and duplications at several loci, such as 7q11.23, 22q11.2, 16p13.3, and 17p11.2.[35] The defects observed in these reciprocal genomic disorders may identically hamper normal cell functions and involve the same organs.[36] Additional experiments, such as comparative transcriptomic studies involving patients with reciprocal genomic disorders, would be required to explore in depth the functional link between gene dosage variations of some genes and the reported phenotype.

Finally, we note that some atypical clinical features were found in only a single patient among the four patients with 1p36.11 duplication (**Table 1**). This could suggest that other gene(s) included in the duplication may participate in the patients' phenotype. Moreover, we cannot rule out a positional effect of the duplication that may disturb other genes located in the vicinity of *ARID1A* and participate in the phenotype.[37,38] Further investigations could be initiated to exclude the implication of other genes outside the MCR or a potential positional effect in the clinical features.

In conclusion, our study shows, for the first time, the clinical and molecular descriptions of a new microduplication syndrome involving the *ARID1A* gene. Although additional work will be necessary to completely elucidate the effects of abnormal *ARID1A* copy-number variations, we provide here the first step to clarifying the pathophysiological mechanism connecting *ARID1A* gene dosage with intellectual disability and microcephaly.

SUPPLEMENTARY MATERIAL
Supplementary material is linked to the online version of the paper at http://www.nature.com/gim

REFERENCES
1. Euskirchen G, Auerbach RK, Snyder M. SWI/SNF chromatin-remodeling factors: multiscale analyses and diverse functions. *J Biol Chem* 2012;287:30897–30905.
2. Brownlee PM, Meisenberg C, Downs JA. The SWI/SNF chromatin remodelling complex: Its role in maintaining genome stability and preventing tumourigenesis. *DNA Repair (Amst)* 2015;32:127–133.
3. Jones S, Li M, Parsons DW, Zhang X, Wesseling J, Kristel P, et al. Coffin-Siris syndrome and related disorders involving components of the BAF (mSWI/SNF) complex: historical review and recent advances using next generation sequencing. *Am J Med Genet C Semin Med Genet* 2014;166C:241–51.
4. Biegel JA, Busse TM, Weissman BE. SWI/SNF chromatin remodeling complexes and cancer. *Am J Med Genet C Semin Med Genet* 2014;166C:350–366.
5. Tsurusaki Y, Okamoto N, Ohashi H, et al. Mutations affecting components of the SWI/SNF complex cause Coffin-Siris syndrome. *Nat Genet* 2012;44:376–378.
6. Coffin GS, Siris E. Mental retardation with absent fifth fingernail and terminal phalanx. *Am J Dis Child* 1970;119:433–439.
7. Kosho T, Miyake N, Carey JC. Coffin-Siris syndrome and related disorders involving components of the BAF (mSWI/SNF) complex: historical review and recent advances using next generation sequencing. *Am J Med Genet C Semin Med Genet* 2014;166C:241–251.

8. Santen GW, Aten E, Sun Y, et al. Mutations in SWI/SNF chromatin remodeling complex gene ARID1B cause Coffin-Siris syndrome. *Nat Genet* 2012;44:379–380.
9. Santen GW, Aten E, Vulto-van Silfhout AT, et al.; Coffin-Siris consortium. Coffin-Siris syndrome and the BAF complex: genotype-phenotype study in 63 patients. *Hum Mutat* 2013;34:1519–1528.
10. Tsurusaki Y, Okamoto N, Ohashi H, Mizuno S, Matsumoto N, Makita Y, et al. Coffin–Siris syndrome is a SWI/SNF complex disorder. *Clin Genet* 2013;85:1–7.
11. Wieczorek D, Bögershausen N, Beleggia F, et al. A comprehensive molecular study on Coffin-Siris and Nicolaides-Baraitser syndromes identifies a broad molecular and clinical spectrum converging on altered chromatin remodeling. *Hum Mol Genet* 2013;22:5121–5135.
12. Mari F, Marozza A, Mencarelli MA, et al. Coffin-Siris and Nicolaides-Baraitser syndromes are a common well recognizable cause of intellectual disability. *Brain Dev* 2015;37:527–536.
13. Son EY, Crabtree GR. The role of BAF (mSWI/SNF) complexes in mammalian neural development. *Am J Med Genet C Semin Med Genet* 2014;166C: 333–349.
14. Coutton C, Bidart M, Rendu J, et al. 190-kb duplication in 1p36.11 including PIGV and ARID1A genes in a girl with intellectual disability and hexadactyly. *Clin Genet* 2013;84:596–599.
15. Romero OA, Sanchez-Cespedes M. The SWI/SNF genetic blockade: effects in cell differentiation, cancer and developmental diseases. *Oncogene* 2014;33: 2681–2689.
16. Krawitz PM, Schweiger MR, Rödelsperger C, et al. Identity-by-descent filtering of exome sequence data identifies PIGV mutations in hyperphosphatasia mental retardation syndrome. *Nat Genet* 2010;42:827–829.
17. Wu RC, Wang TL, Shih IeM. The emerging roles of ARID1A in tumor suppression. *Cancer Biol Ther* 2014;15:655–664.
18. Bidart M, Coutton C. Intellectual disability, oncogenes and tumour suppressor genes: the way forward? *J Genet* 2012;91:257–258.
19. Santen GW, Kriek M, van Attikum H. SWI/SNF complex in disorder: SWItching from malignancies to intellectual disability. *Epigenetics* 2012;7:1219–1224.
20. Shigetomi H, Oonogi A, Tsunemi T, et al. The role of components of the chromatin modification machinery in carcinogenesis of clear cell carcinoma of the ovary (review). *Oncol Lett* 2011;2:591–597.
21. Zeng Y, Liu Z, Yang J, Liu Y, Huo L, Li Z, et al. ARID1A is a tumor suppressor and inhibits glioma cell proliferation via the PI3K pathway. *Head Neck Oncol* 2013;5:6.
22. Lessard J, Wu JI, Ranish JA, et al. An essential switch in subunit composition of a chromatin remodeling complex during neural development. *Neuron* 2007;55:201–215.
23. Rakic P. A small step for the cell, a giant leap for mankind: a hypothesis of neocortical expansion during evolution. *Trends Neurosci* 1995;18:383–388.
24. Barbelanne M, Tsang WY. Molecular and cellular basis of autosomal recessive primary microcephaly. *Biomed Res Int* 2014;2014:547986.
25. Bicknell LS, Walker S, Klingseisen A, et al. Mutations in ORC1, encoding the largest subunit of the origin recognition complex, cause microcephalic primordial dwarfism resembling Meier-Gorlin syndrome. *Nat Genet* 2011;43:350–355.
26. Ostergaard P, Simpson MA, Mendola A, et al. Mutations in KIF11 cause autosomal-dominant microcephaly variably associated with congenital lymphedema and chorioretinopathy. *Am J Hum Genet* 2012;90:356–362.
27. Morris-Rosendahl DJ, Kaindl AM. What next-generation sequencing (NGS) technology has enabled us to learn about primary autosomal recessive microcephaly (MCPH). *Mol Cell Probes* 2015;29:271–281.
28. Woods CG. Human microcephaly. *Curr Opin Neurobiol* 2004;14:112–117.
29. Marthiens V, Rujano MA, Pennetier C, Tessier S, Paul-Gilloteaux P, Basto R. Centrosome amplification causes microcephaly. *Nat Cell Biol* 2013;15: 731–740.
30. Martin CA, Ahmad I, Klingseisen A, et al. Mutations in PLK4, encoding a master regulator of centriole biogenesis, cause microcephaly, growth failure and retinopathy. *Nat Genet* 2014;46:1283–1292.
31. Shaheen R, Al Tala S, Almoisheer A, Alkuraya FS. Mutation in PLK4, encoding a master regulator of centriole formation, defines a novel locus for primordial dwarfism. *J Med Genet* 2014;51:814–816.
32. Kishi K, van Vugt MA, Okamoto K, Hayashi Y, Yaffe MB. Functional dynamics of Polo-like kinase 1 at the centrosome. *Mol Cell Biol* 2009;29:3134–3150.
33. Sakai D, Dixon J, Dixon MJ, Trainor PA. Mammalian neurogenesis requires Treacle-Plk1 for precise control of spindle orientation, mitotic progression, and maintenance of neural progenitor cells. *PLoS Genet* 2012;8:e1002566.
34. Barber JC, Hall V, Maloney VK, et al. 16p11.2-p12.2 duplication syndrome; a genomic condition differentiated from euchromatic variation of 16p11.2. *Eur J Hum Genet* 2013;21:182–189.
35. Tassano E, Alpigiani MG, Calcagno A, et al. Clinical and molecular delineation of a 16p13.2p13.13 microduplication. *Eur J Med Genet* 2015;58:194–198.
36. Pérez de Castro I, de Cárcer G, Malumbres M. A census of mitotic cancer genes: new insights into tumor cell biology and cancer therapy. *Carcinogenesis* 2007;28:899–912.
37. Kleinjan DJ, van Heyningen V. Position effect in human genetic disease. *Hum Mol Genet* 1998;7:1611–1618.
38. Blumenthal I, Ragavendran A, Erdin S, et al. Transcriptional consequences of 16p11.2 deletion and duplication in mouse cortex and multiplex autism families. *Am J Hum Genet* 2014;94:870–883.