

# Exon-level array CGH in a large clinical cohort demonstrates increased sensitivity of diagnostic testing for Mendelian disorders

Swaroop Aradhya, PhD<sup>1</sup>, Rachel Lewis, MS<sup>1</sup>, Tahrra Bonaga, BS<sup>1</sup>, Nnenna Nwokekeh, MS<sup>1</sup>, Amanda Stafford, MS<sup>1</sup>, Barbara Boggs, PhD<sup>1</sup>, Kathleen Hruska, PhD<sup>1</sup>, Nizar Smaoui, MD<sup>1</sup>, John G. Compton, PhD<sup>1</sup>, Gabriele Richard, MD<sup>1</sup> and Sharon Suchy, PhD<sup>1</sup>

**Purpose:** Mendelian disorders are most commonly caused by mutations identifiable by DNA sequencing. Exonic deletions and duplications can go undetected by sequencing, and their frequency in most Mendelian disorders is unknown.

**Methods:** We designed an array comparative genomic hybridization (CGH) test with probes in exonic regions of 589 genes. Targeted testing was performed for 219 genes in 3,018 patients. We demonstrate for the first time the utility of exon-level array CGH in a large clinical cohort by testing for 136 autosomal dominant, 53 autosomal recessive, and 30 X-linked disorders.

**Results:** Overall, 98 deletions and two duplications were identified in 53 genes, corresponding to a detection rate of 3.3%. Approximately 40% of positive findings were deletions of only one or two exons. A high frequency of deletions was observed for several auto-

somal dominant disorders, with a detection rate of 2.9%. For autosomal recessive disorders, array CGH was usually performed after a single mutation was identified by sequencing. Among 138 individuals tested for recessive disorders, 10.1% had intragenic deletions. For X-linked disorders, 3.5% of 313 patients carried a deletion or duplication.

**Conclusion:** Our results demonstrate that exon-level array CGH provides a robust option for intragenic copy number analysis and should routinely supplement sequence analysis for Mendelian disorders.

*Genet Med* 2012;14(6):594–603

**Key words:** array CGH; copy number; deletion; duplication; exon; Mendelian

## INTRODUCTION

The majority of mutations in Mendelian disorders are detected by DNA sequencing. However, in recent years the significance and frequency of pathogenic intragenic deletions or duplications (copy number mutations) have become increasingly evident. Quantitative methods are now used for molecular diagnosis of a limited number of disorders.<sup>1</sup> In a variety of genetic disorders, including Rett syndrome, Smith–Magenis syndrome, and Prader–Willi syndrome, routine testing includes copy number analysis by multiplex ligation-dependent amplification (MLPA) assay, quantitative PCR (qPCR), or fluorescence in situ hybridization. The same methods have also been used to identify larger chromosomal rearrangements, including those affecting subtelomeric regions.<sup>2,3</sup> Array comparative genomic hybridization (CGH) has proven to be a powerful tool for copy number analysis but has been used mostly for cytogenetic analysis to detect large genomic deletions and duplications that extend hundreds of kilobases to megabases.<sup>4</sup> Some recent reports have shown that high-resolution array CGH with probes densely distributed across individual genes can detect small deletions or duplications, but this approach has not been widely applied in a

clinical setting.<sup>5,6</sup> We constructed a custom high-density oligonucleotide array with probes targeted to the individual exons of 589 genes associated with known genetic disorders to identify whole and partial gene deletions or duplications. Our results from testing 3,018 patients with this array demonstrate that exon array CGH complements DNA sequencing and increases the mutation detection rate in the molecular diagnosis of autosomal and X-linked Mendelian disorders.

## MATERIALS AND METHODS

### Array design

The 3,018 clinical cases tested by exon array CGH were analyzed on one of the two array versions. On the first array version, 1,584 cases were tested; 1,434 cases were tested on the second. The first version included two or more probes in most exons of 465 targeted genes and three probes in each intron, regardless of size. Coverage of some small exons included probes in the intronic sequence immediately flanking the exon, based on the premise that any deletion/duplication of that exon would probably extend into the intron. Genes with unprocessed pseudogenes were not included in the target list. The second version

<sup>1</sup>GeneDx, Gaithersburg, Maryland, USA. Correspondence: Swaroop Aradhya ([swaroop@genedx.com](mailto:swaroop@genedx.com))

Submitted 21 September 2011; accepted 28 November 2011; advance online publication 1 March 2012. doi:[10.1038/gim.2011.65](https://doi.org/10.1038/gim.2011.65)

of the exon array included seven probes across each exon and the flanking 250-bp intronic sequence on either side. The array was designed to include probe coverage of 589 genes, including all 465 genes from the first array version. Intronic regions other than the 250-bp exonic flanks were not covered, and deletions or duplications in those regions would not be detected. Genes with processed pseudogenes (e.g., *PTEN*) were furnished with extra probe coverage within the 250-bp exon-flanking intronic sequences. Untranslated regions of the exons were also covered if part of those exons included coding sequences. The average gene size and the exon size in this collection of 589 genes were 104 kb and 741 bp, respectively. Each array version was validated with 60–80 DNA samples that previously showed either normal copy number or a known MLPA- or qPCR-confirmed deletion or duplication in one of a variety of genes (data not shown).

Following validation of the exon array, prospective studies were performed on blood samples from patients referred to our clinical laboratory for deletion/duplication testing for various Mendelian disorders. For autosomal recessive disorders, deletion/duplication testing was performed if DNA sequencing identified only a single mutation. For individuals with an autosomal dominant disorder or for female carriers of an X-linked disorder (affected or unaffected), deletion/duplication testing was either offered as a second test if gene sequencing was negative or in combination with the sequencing test when that disorder had a known high frequency of deletions.

#### Array hybridization and data analysis

DNA was extracted from blood samples on the QiaCube (Qiagen, Valencia, CA) automated system. Labeling was carried out with the Enzo CGH labeling kit for oligo arrays (Enzo Life Sciences, Plymouth Meeting, PA). Array CGH was performed with 0.5 µg of DNA according to the manufacturer's protocol (Agilent Technologies, Santa Clara, CA). The data for each patient were examined only for the specific gene (or gene panel) requested. Data were analyzed using the ADM-1 algorithm in DNA Analytics/Genomic Workbench software. Reportable data were based on  $\log_2$  ratio deviations >0.25 and including two or more adjacent probes.

#### Confirmation of array findings

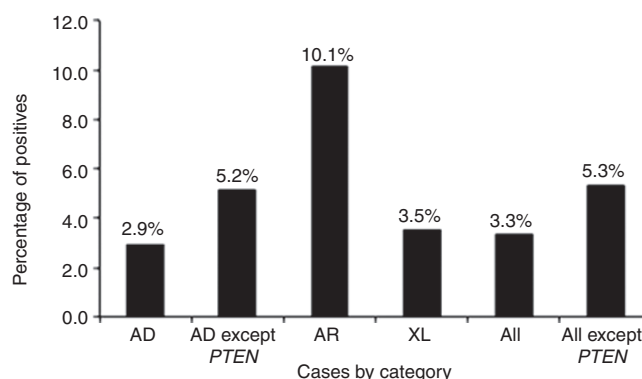
Probe deviations including two or more adjacent probes were confirmed by qPCR with custom-designed primers, MLPA or, in one case, whole-genome array CGH. MLPA was performed according to the manufacturer's protocols (MRC-Holland, Amsterdam, The Netherlands). The following SALSA MLPA kits were validated for clinical use: P067 PTCH1, kit P187 Holoprosencephaly HPE, P225-B1 PTEN, P101 STK11, P219-B1 PAX6, P313 CREBBP, P215 EXT, P180 Limb Malformations-2, and P015 MECP2. qPCR analysis was performed using a TaqMan assay based on the Human Universal Probe Library set from Roche Applied Science (Roche, Indianapolis, IN; <https://www.roche-applied-science.com/sis/rtpcr/upl/index.jsp?id=UP030000>). Custom primers were designed using the

online Universal Probe Library assay design center (<https://www.roche-applied-science.com>). Seventy nanograms each of patient sample and control DNA samples were added to separate reaction mixtures containing a TaqMan probe and FastStart TaqMan mastermix (Roche) and locus-specific primers. The amplification was carried out at 60°C annealing temperature for 45 cycles on a Stratagene Mx3000P/3005P machine (Agilent Technologies). Three normal genomes were tested in triplicate along with the patient sample. In a valid positive assay, the  $C_t$  for the clinical sample showed a difference of at least one PCR cycle relative to the normal sample, the triplicates deviated by <0.30  $C_t$ , and all negative controls showed normal copy number for the locus tested. The assay used an internal normalizing target (*SOD1* gene in 21q22.11) to ensure that equal amounts of DNA were used in all tested samples. The difference in  $C_t$  values between the clinical sample and one normal control sample was expressed as a fold change in the copy number of the target gene.

## RESULTS

### Clinical testing of 3,018 cases

Exon array CGH was used to examine 3,018 patients for deletions or duplications in 219 genes. More than one gene was analyzed in 307 individuals affected with a genetically heterogeneous disorder. Therefore, a total of 4,354 genes were analyzed in the 3,018 individuals. Exon array CGH identified 98 partial or whole-gene deletions and two duplications, corresponding to a detection rate of 3.3% in the individuals tested (Figure 1). qPCR, MLPA, or whole-genome array CGH confirmed copy number mutations detected by exon array CGH. Table 1 and Supplementary Table S5 online lists all the copy number mutations identified in our cohort, and Figure 2 shows selected examples of partial gene deletions or duplications detected by exon array CGH. Copy number mutations were identified in 53 of the 219 genes tested. No deletions or duplications were found in the remaining 166 genes, although most of these were evaluated in 10 cases or fewer. Forty percent of the 3,018 cases



**Figure 1 Percentage of exon array CGH results classified according to mode of inheritance in 100 positive cases.** Results are shown with and without the *PTEN* cases included. AD, dominant; all, all cases tested by exon array CGH; AR, recessive; CGH, comparative genomic hybridization; XL, X-linked.

**Table 1** Copy number mutations identified by exon array CGH in 3,018 patients

Gene	Disorder	Inheritance	Mutation	Confirmation
<i>ALMS1</i>	Alstrom syndrome	AR	Het del exon 13–16	qPCR
<i>BTK</i>	Agammaglobulinemia	XL	Nullizygous del exon 6–7	qPCR
<i>CBS</i>	Homocystinuria	AR	Het del exon 11–12	qPCR
<i>COL1A1</i>	Osteogenesis imperfecta, type II	AD	Het del whole gene	qPCR
<i>COL7A1</i>	Epidermolysis bullosa dystrophica	AR	Het del exon 1–2	qPCR
<i>CREBBP</i>	Rubinstein–Taybi syndrome	AD	Het del exon 17–21	MLPA
<i>CREBBP</i>	Rubinstein–Taybi syndrome	AD	Het del exon 1–26	MLPA
<i>CREBBP</i>	Rubinstein–Taybi syndrome	AD	Het del exon 16	MLPA
<i>CREBBP</i>	Rubinstein–Taybi syndrome	AD	Het del whole gene	MLPA
<i>CREBBP</i>	Rubinstein–Taybi syndrome	AD	Het del exon 4–6	MLPA
<i>CREBBP</i>	Rubinstein–Taybi syndrome	AD	Het del exon 2	MLPA
<i>CREBBP</i>	Rubinstein–Taybi syndrome	AD	Het del whole gene	MLPA
<i>CTNS</i>	Nephropathic cystinosis	AR	Hom del exon 1–3	qPCR
<i>DCLRE1C</i>	Severe combined immunodeficiency	AR	Hom del exon 1–3	qPCR
<i>DHCR7</i>	Smith–Lemli–Opitz syndrome	AR	Het del exon 3–4	qPCR
<i>EDA1</i>	Ectodermal dysplasia	XL	Het del exon 2	qPCR
<i>EDA1</i>	Ectodermal dysplasia	XL	Het del exon 1	qPCR
<i>EXT1</i>	Multiple exostoses, type 1	AD	Het del exon 1	MLPA
<i>EXT1</i>	Multiple exostoses, type 1	AD	Het del whole gene	MLPA
<i>EXT2</i>	Multiple exostoses, type II	AD	Het del whole gene	MLPA
<i>EXT2</i>	Multiple exostoses, type II	AD	Het del exon 12	MLPA
<i>EYA1</i>	Branchiootorenal syndrome	AD	Het del exon 1–12	qPCR
<i>EYA1</i>	Branchiootorenal syndrome	AD	Het del whole gene	qPCR
<i>FH</i>	Leiomyomatosis and renal cell cancer	AD	Het del exon 6	MLPA
<i>FLCN</i>	Birt–Hogg–Dube syndrome	AD	Het del exon 1	qPCR
<i>FOXC1</i>	Axenfeld–Rieger syndrome	AD	Het del whole gene	qPCR
<i>FOXC1</i>	Axenfeld–Rieger syndrome	AD	Het del whole gene	qPCR
<i>GAA</i>	Glycogen storage disease II	AR	Het del exon 18	qPCR
<i>GHR</i>	Laron syndrome	AR	Hom del exon 7–10	qPCR
<i>HADHB</i>	Trifunctional protein deficiency	AR	Het del exon 4	qPCR
<i>IRF6</i>	Van der Woude syndrome	AD	Het del whole gene	qPCR
<i>JAG1</i>	Alagille syndrome	AD	Het del exon 3	MLPA
<i>JAG1</i>	Alagille syndrome	AD	Het del whole gene	MLPA
<i>JAG1</i>	Alagille syndrome	AD	Het del exon 6–26	MLPA
<i>KCNQ1</i>	Long QT syndrome 1	AD	Het del exon 7–10	qPCR
<i>KCNQ1</i>	Long QT syndrome 1	AD	Het del exon 3	qPCR
<i>MECP2</i>	Rett syndrome	XL	Het del exon 3–4	MLPA
<i>MECP2</i>	Rett syndrome	XL	Het dup exon 3–4	MLPA
<i>MEN1</i>	Multiple endocrine neoplasia, type 1	AD	Het del whole gene	MLPA
<i>MOCS2</i>	Molybdenum cofactor deficiency	AR	Het del exon 3–5	qPCR
<i>MYCN</i>	Feingold syndrome	AD	Het del whole gene	qPCR

See **Supplementary Table S5** online for genomic coordinates and transcript information.

AD, autosomal dominant; AR, autosomal recessive; del, deletion; dup, duplication; Het, heterozygous; Hom, homozygous; MLPA, multiplex ligation-dependent probe amplification; qPCR, quantitative polymerase chain reaction; XL, X-linked.

**Table 1** Continued on next page

Table 1 Continued.

Gene	Disorder	Inheritance	Mutation	Confirmation
<i>NDP</i>	Norrie disease	XL	Nullizygous del whole gene	qPCR
<i>NSD1</i>	Sotos syndrome	AD	Het del exon 5	qPCR
<i>OTC</i>	Ornithine transcarbamylase deficiency	XL	Het del whole gene	qPCR
<i>OTX2</i>	Microphthalmia	AD	Het del whole gene	qPCR
<i>PAH</i>	Phenylketonuria	AR	Het del exon 5	qPCR
<i>PAX6</i>	Aniridia	AD	Mosaic del exon 8–12	MLPA
<i>PAX6</i>	Aniridia	AD	Het del exon 6–7	MLPA
<i>PAX6</i>	Aniridia	AD	Het del exon 1–7	MLPA
<i>PAX6 and ELP4</i>	Aniridia	AD	Het del ELP4 exon 10–PAX6 exon 7–13	MLPA
<i>PAX6 and ELP4</i>	Aniridia	AD	Het del ELP4 exon 10–PAX6 exon 13	MLPA
<i>PAX6 and ELP4</i>	Aniridia	AD	Het del ELP4 exon 5–10, PAX6 whole gene del	MLPA
<i>PAX6 and ELP4</i>	Aniridia	AD	Het del PAX6 exon 8–13 and ELP4 exon 10	MLPA
<i>PCCA</i>	Propionic academia	AR	Het del exon 20–21	qPCR
<i>PORCN</i>	Focal dermal hypoplasia	XL	Het del whole gene	qPCR
<i>PTCH1</i>	Gorlin syndrome	AD	Het del exon 19–22	MLPA
<i>PTCH1</i>	Gorlin syndrome	AD	Het del exon 16–24	MLPA
<i>PTCH1</i>	Gorlin syndrome	AD	Het del exon 21–24	MLPA
<i>PTCH1</i>	Gorlin syndrome	AD	Het del whole gene	MLPA
<i>PTCH1</i>	Gorlin syndrome	AD	Mosaic del exon 13–15	MLPA
<i>PTCH1</i>	Gorlin syndrome	AD	Het del whole gene	MLPA
<i>PTCH1</i>	Gorlin syndrome	AD	Mosaic del exon 6–7	MLPA
<i>PTEN</i>	Cowden syndrome	AD	Het dup exon 1–5	MLPA
<i>PTEN</i>	Cowden syndrome	AD	Het del whole gene	MLPA
<i>PTEN</i>	Cowden syndrome	AD	Het del whole gene	MLPA
<i>PTEN</i>	Cowden syndrome	AD	Het del exon 3–9	MLPA
<i>PTEN</i>	Cowden syndrome	AD	Het del whole gene	MLPA
<i>PTEN</i>	Cowden syndrome	AD	Mosaic del exon 6–9	MLPA
<i>RAI1</i>	Smith–Magenis syndrome	AD	Het del whole gene	qPCR
<i>RPS6KA3</i>	Coffin–Lowry syndrome	XL	Het del exon 10	MLPA
<i>RS1</i>	X-linked juvenile retinoschisis	XL	Het del exon 2	Southern blot (performed elsewhere) <sup>38</sup>
<i>SALL4</i>	Duane-radial ray syndrome	AD	Het del whole gene	MLPA
<i>SDHB</i>	Paragangliomas 4	AD	Het del exon 3	MLPA
<i>SDHB</i>	Paragangliomas 4	AD	Het del exon 3	MLPA
<i>SDHD</i>	Paragangliomas 1	AD	Het del exon 4	MLPA
<i>SERPING1</i>	Hereditary angioedema	AD	Het del exon 1–6	qPCR
<i>SHANK3</i>	Chromosome 22q13.3 deletion syndrome	AD	Het del whole gene	Array CGH
<i>SOX2</i>	Microphthalmia	AD	Het whole gene del	MLPA
<i>SOX2</i>	Microphthalmia	AD	Het del whole gene	qPCR
<i>SOX2</i>	Microphthalmia	AD	Het del whole gene	qPCR

See **Supplementary Table S5** online for genomic coordinates and transcript information.

AD, autosomal dominant; AR, autosomal recessive; del, deletion; dup, duplication; Het, heterozygous; Hom, homozygous; MLPA, multiplex ligation-dependent probe amplification; qPCR, quantitative polymerase chain reaction; XL, X-linked.

Table 1 Continued on next page

Table 1 Continued.

Gene	Disorder	Inheritance	Mutation	Confirmation
<i>SOX9</i>	Campomelic dysplasia	AD	Het del whole gene	qPCR
<i>STK11</i>	Peutz–Jeghers syndrome	AD	Het del exon 2–3	MLPA
<i>STK11</i>	Peutz–Jeghers syndrome	AD	Het del whole gene	MLPA
<i>STK11</i>	Peutz–Jeghers syndrome	AD	Het del exon 6–8	MLPA
<i>STK11</i>	Peutz–Jeghers syndrome	AD	Mosaic del exon 6–10	MLPA
<i>STK11</i>	Peutz–Jeghers syndrome	AD	Het del whole gene	MLPA
<i>STK11</i>	Peutz–Jeghers syndrome	AD	Het del exon 2–3	MLPA
<i>STK11</i>	Peutz–Jeghers syndrome	AD	Het del exon 1	MLPA
<i>STK11</i>	Peutz–Jeghers syndrome	AD	Het del exon 1	MLPA
<i>STK11</i>	Peutz–Jeghers syndrome	AD	Het del exon 1	MLPA
<i>STK11</i>	Peutz–Jeghers syndrome	AD	Het del exon 1	MLPA
<i>TAZ</i>	Dilated cardiomyopathy	XL	Nullizygous del exon 1–5	qPCR
<i>TBX5</i>	Holt–Oram syndrome	AD	Het del whole gene	MLPA
<i>TP63</i>	Ectrodactyly, ectodermal dysplasia, and cleft lip/palate syndrome 3	AD	Het del exon 1–4	qPCR
<i>TWIST1</i>	Saethre–Chotzen syndrome	AD	Het del whole gene	qPCR
<i>VHL</i>	Von Hippel–Lindau syndrome	AD	Het del whole gene	qPCR
<i>VHL</i>	Von Hippel–Lindau syndrome	AD	Het del exon 1–2	qPCR
<i>VHL</i>	Von Hippel–Lindau syndrome	AD	Het del exon 3	qPCR
<i>VPS13B</i>	Cohen syndrome	AR	Het del exon 34	qPCR
<i>VPS13B</i>	Cohen syndrome	AR	Het del exon 31	qPCR

See **Supplementary Table S5** online for genomic coordinates and transcript information.

AD, autosomal dominant; AR, autosomal recessive; del, deletion; dup, duplication; Het, heterozygous; Hom, homozygous; MLPA, multiplex ligation-dependent probe amplification; qPCR, quantitative polymerase chain reaction; XL, X-linked.

were referred for *PTEN* testing. The overall detection rate for copy number mutations by exon array CGH for Mendelian disorders in our cohort excluding *PTEN* cases was 5.3% (95/1,787; **Figure 1**).

### Autosomal dominant disorders

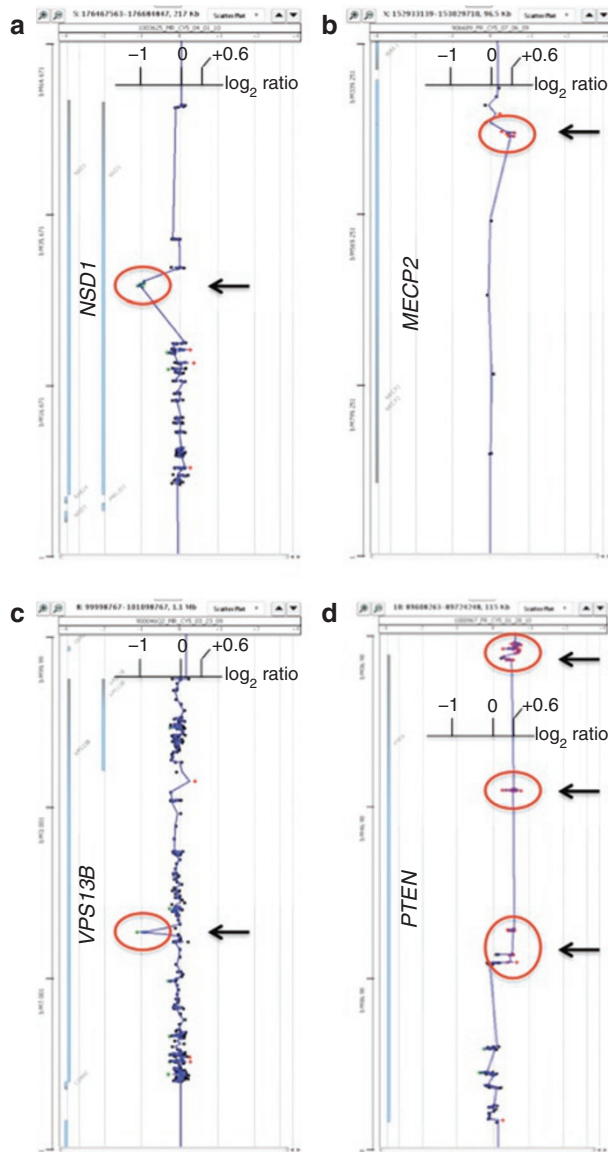
Patients tested for autosomal dominant disorders, most of which were caused by haploinsufficiency, constituted a group of 2,567 individuals. Almost half of these individuals ( $n = 1,231$ ) were specifically referred for *PTEN* testing. The remaining 1,336 individuals diagnosed with other autosomal dominant disorders were tested for copy number mutations in one or more of a variety of genes. For 13 of those genes, array testing was performed concurrently with sequencing, while other genes were tested after a mutation was excluded by sequencing. Sixty-nine of the 1,336 individuals carried a gene deletion, yielding a positive rate of 5.2% for this group of dominant disorders (**Figure 1**). The majority of individuals who tested positive had partial gene deletions involving one or more exons; whole-gene deletions accounted for 34% of all mutations observed.

The frequency of whole or partial deletions and duplications in patients sent for *PTEN* analysis with test indications of Bannayan–Riley–Ruvalcaba syndrome (BRRS), Cowden syndrome (CS), macrocephaly, and/or autism spectrum disorder was 0.5%, which is lower than the previously published

rate seen in smaller cohorts of patients with normal results by sequencing.<sup>7,8</sup> Whole-gene deletions in *PTEN* were observed in three patients with BRRS, CS, or Sotos syndrome–like overgrowth. One patient with developmental delay and multiple lipomas had a partial gene deletion including exons 3–9, and two patients with suspected CS showed mosaicism for a partial deletion of exons 6–9 and a partial duplication of *PTEN* involving exons 1–5, respectively (**Figure 2d**). No copy number mutations were observed in individuals referred for macrocephaly/microcephaly associated with autism spectrum disorders.

In agreement with published literature, a high rate of pathogenic gene deletions was identified in several autosomal dominant disorders in our data set (see **Supplementary Table S1** online). For Peutz–Jeghers syndrome, 10 out of 31 patients (32%) with negative sequencing results carried a deletion detected by exon array CGH in the *STK11* gene. This is comparable to previously published reports of 16–30% incidence of exonic deletions in *STK11*.<sup>9</sup> Two whole-gene and eight partial *STK11* deletions were observed, including four cases with an apparently recurrent deletion of exon 1. The positive rate of 9.3% for *CREBBP* intragenic deletions in 75 patients with Rubinstein–Taybi syndrome and of 7.3% for *EXT1* or *EXT2* deletions in 55 patients with multiple exostoses falls in the range reported by others.<sup>10</sup> Finally, consistent with previous reports, 1 out of 13 patients





**Figure 2** Examples of exon array comparative genomic hybridization data from four separate genes. Arrows mark the probes that identified the deletion or duplication. (a) A heterozygous deletion of exon 5 in *NSD1* in a patient with Sotos syndrome. (b) A partial gene duplication including exons 3–4 in *MECP2* in a female patient with Rett syndrome. (c) A heterozygous deletion of exon 31 in *VPS13B* (COH1) in a patient with Cohen syndrome; this patient also had a sequence change on the other allele. (d) A partial gene duplication including exons 1–5 in *PTEN* in a patient with Cowden syndrome.

(8%) showed a deletion in the *MYCN* gene in individuals with Feingold syndrome, and 1 out of 11 patients (9%) had a deletion in the *SOX9* gene associated with campomelic dysplasia.<sup>11</sup>

We also identified seven deletions in the *PTCH1* gene, including two mosaic mutations, in 128 patients diagnosed with Gorlin syndrome (basal cell nevus syndrome). The incidence of *PTCH1* deletions was, therefore, over 5%, and most of these were private mutations. This represents the largest number of patients tested concurrently for sequencing and copy number mutations in *PTCH1*.<sup>11,12</sup>

For some disorders the positive rate in our data set was lower than that published elsewhere (see **Supplementary Table S1** online). For example, only 3 out of 116 patients (2.6%) with von Hippel–Lindau syndrome had a deletion in the *VHL* gene, while the published deletion rate is up to 30%.<sup>13</sup> This discrepancy is probably attributable to the fact that molecular testing is requested after a gene deletion has been excluded by fluorescence in situ hybridization. Likewise, only 3 out of 204 (1.5%) patients with Alagille syndrome tested by sequencing and exon array CGH had a gene deletion in our cohort, whereas fluorescence in situ hybridization analysis elsewhere has revealed gross deletions, including *JAG1*, in 5.7% of patients.<sup>14</sup> Other disorders with lower than expected deletion rate in our data set were hereditary angioedema (*SERPING1*: 1/27 or 3.7%), branchiootorenal syndrome (*EYA1*: 2/41 or 4.9%), and Duane radial ray syndrome (*SALL4*: 1/58 or 1.7%; **Table 1**).<sup>15,16</sup>

We also identified unique deletions in several other dominant disorders after DNA sequencing failed to reveal a pathogenic mutation (**Table 1**). Testing for developmental eye disorders revealed four exonic deletions involving the *SOX2* and *OTX2* genes in 26 (15.4%) individuals with anophthalmia or microphthalmia and two *FOXC1* deletions in 12 (17%) patients with Axenfeld–Rieger syndrome. Likewise, 7 out of 36 patients (19%) with a clinical indication of aniridia had a deletion in the *PAX6* gene and previous reports show a wide range for the incidence of *PAX6* deletions.<sup>17,18</sup> Notably, more than half of the *PAX6* deletions that we found extended into the neighboring *ELP4* gene but did not include the *WT1* gene. Deletions were also detected in individuals with Birt–Hogg–Dube syndrome (*FLCN*: 1/15 or 6.7%), Holt–Oram syndrome (*TBX5*: 1/21 or 4.5%), long QT syndrome (*KCNQ1*: 2/71 or 2.8%), and multiple endocrine neoplasia (*MEN1*: 1/43 or 2.3%; **Table 1**). Pathogenic deletions in the disorders above are rare and deletion testing is not routinely performed. This applies to many other disorders as well. For instance, although only a small number (one to three) of individuals were referred for testing, we identified novel pathogenic intragenic deletions in *IRF6* (Van der Woude syndrome), *NSD1* (Sotos syndrome), *SHANK3* (22q13.3 subtelomeric deletion syndrome), *TWIST1* (Saethre–Chotzen syndrome), and *TP63* (ectrodactyly ectodermal dysplasia and cleft lip/palate syndrome 3; **Table 1**).

### Autosomal recessive disorders

We tested 138 individuals by exon array CGH analysis for autosomal recessive disorders and identified 14 deletions of one or more exons (see **Supplementary Table S2** online). No duplications were identified. Fifty-four cases were referred with an indication of an inborn error of metabolism (IEM). Seven of these cases were positive for a deletion and represented a positive rate of 13.0%, compared with 8.3% for other indications (see **Supplementary Table S3** online).

Most of the 138 cases had sequencing prior to or concurrently with deletion/duplication testing. Of 93 individuals with a single pathogenic mutation identified by sequencing, 10 were also found to have a whole or partial gene deletion, which corresponded to

a positive rate of 10.8% (see **Supplementary Table S3** online). In 45 IEM cases with one mutation detected by sequencing, we identified seven deletions (15.6%). Of the 48 cases referred for other indications, only three were positive (6.3%). Forty-five individuals who had no prior sequencing or with negative sequencing results had deletion/duplication testing, and four (10.1%) were homozygous for intragenic deletions.

### X-linked disorders

Exon array CGH was performed in 313 individuals for deletions or duplications in X-linked genes. Eleven patients showed a pathogenic copy number mutation, representing a positive rate of 3.5% (see **Supplementary Table S4** online and **Figure 1**). One patient had partial gene duplication and the remainder carried a whole or partial gene deletion. Eight of the 11 individuals with a positive exon array CGH result were females diagnosed with one of several conditions (hypohidrotic ectodermal dysplasia, focal dermal hypoplasia, juvenile retinoschisis, Rett syndrome, Coffin–Lowry syndrome, and ornithine transcarbamylase deficiency). The three remaining deletions were found in males with agammaglobulinemia, dilated cardiomyopathy, or Norrie disease.

### Detection of mosaicism

Array CGH has limited sensitivity to detect mosaic genomic changes. Mosaicism for  $\geq 25\%$  mutant allele is reliably detectable using whole-genome cytogenetic arrays.<sup>4</sup> Using the exon array, we identified five mosaic mutations in four genes: one each in *PTEN*, *STK11*, and *PAX6* and two in *PTCH1* (**Table 1**

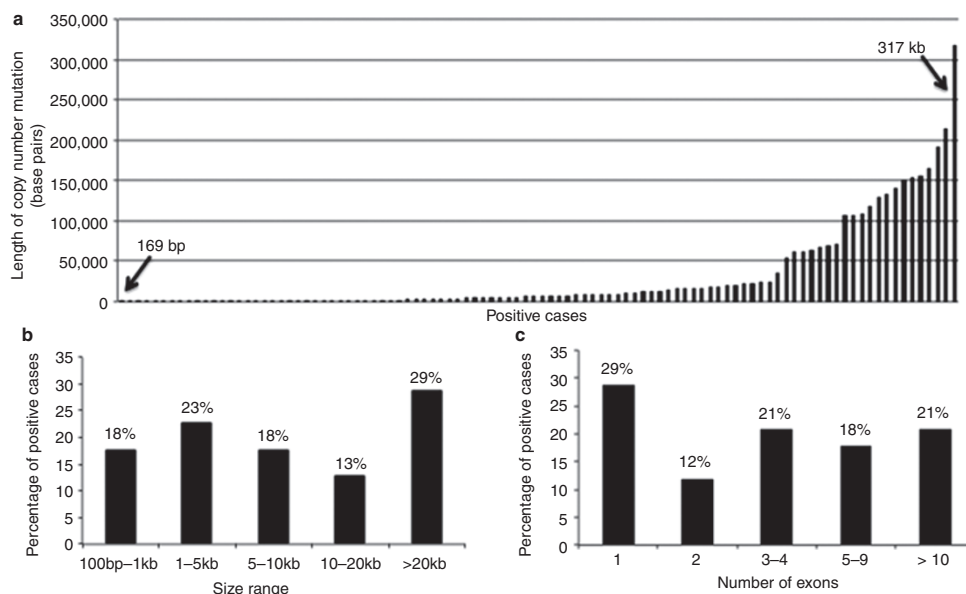
and see **Supplementary Figure S1** online). In each case, the patients with the mosaic deletions were clinically affected with the disorders associated with those specific genes. MLPA was used to confirm the mosaic deletion in all five instances, although the level of mosaicism could not be accurately determined.

### Size of intragenic copy number changes

The minimum size of the detected exonic copy number mutations ranged from ~169 bp to 317 kb (**Figure 3a**). Forty-one percent of the deletions and duplications were  $< 5$  kb, and 42% included only one or two exons (**Figure 3b,c**). A deletion as small as 169 bp was detected by four 60-mer probes that had significant sequence overlap. The larger deletions and duplications included complete genes. For all positive cases, the minimum size of a deletion/duplication was determined by the sequence positions of the probes that showed abnormal copy number. The break points are located at an undetermined site within regions flanking neighboring probes that showed normal copy number. Because our targeted exon-centered array did not contain a whole-genome backbone, the size of deletions/duplications extending beyond the 5'- and 3'-ends of genes could not be determined.

## DISCUSSION

In this study, we evaluated 3,018 patients with a suspected Mendelian disorder for intragenic deletions/duplications by exon-level array CGH in one or more of 219 genes and found 100 mutations (2.9%). Excluding the large number of cases tested



**Figure 3** Minimum size distribution of deletions and duplications identified by exon array comparative genomic hybridization. (a) Sizes of all deletions and duplications detected in 100 positive cases. The smallest deletion was 169 bp and the largest deletion was a whole-gene deletion of 317 kb. (b) Size of deletions and duplications by ranges. Forty-one percent of mutations were below 5 kb. (c) Deletions and duplications separated by number of exons affected. Note the high proportion of single- and dual-exon deletions.

for a single gene (*PTEN*) with a low frequency of deletions, the rate of deletion/duplications in the remaining genes was 5.3%. Whole-gene deletions accounted for 34% of mutations. Most of the remaining partial copy number mutations were private. An important observation from our data is that 41% of the exonic copy number mutations were <5 kb and/or involved only one or two exons. There are no previous reports describing a large number of copy number mutations and showing the distribution based on size or the number of exons affected. This is the first report of exon-level array CGH testing of a very large clinical cohort. Even with our exon-targeted array design that does not interrogate promoters or introns, our data demonstrate that intragenic copy number mutations are more prevalent than perhaps previously suspected in Mendelian disorders as a whole, and should be routinely considered in the diagnostic workup of these disorders. Exon array CGH is a robust, high-resolution copy number assay that fits this purpose.

Ninety-eight percent of mutations we identified were deletions. A variety of other studies have also reported a higher proportion of deletions among copy number mutations within single genes.<sup>19,20</sup> Several explanations may account for this observation. For example, carrying a complete extra copy of a gene is often not pathogenic.<sup>21</sup> Another potential explanation is that duplications including the first or last exon of a gene may leave the final gene structure sufficiently preserved for normal transcription. For these reasons, individuals with deletions are more likely to manifest a phenotype and seek medical care. Finally, as with large intrachromosomal rearrangements, the mechanism by which intragenic copy number mutations occur may favor the formation of deletions over duplications. Because most of the mutations in our data appear to be nonrecurrent, the mechanism for their formation is probably not based on homologous recombination but rather on one of a variety of other mechanisms.<sup>22</sup>

We observed a high frequency of pathogenic intragenic copy number mutations in autosomal recessive disorders when a single heterozygous mutation was identified by sequencing (10.8%). An even higher rate was observed for IEMs, with a 15.6% detection rate in cases in which a single mutation was identified by sequencing. This high rate of deletions detected for IEMs is not surprising because molecular diagnosis is usually performed after positive biochemical results indicate a likely candidate gene. The identification of both disease alleles is important to confirm a diagnosis, for family counseling, and for carrier and prenatal genetic testing.

To our knowledge, this is the first case of an intragenic exon-level deletion in the *DHCR7*,<sup>23</sup> *HADHB*, and *MOCS2* genes. There have been previous reports of exon-level deletions in other genes reported here, including *GAA*, *PAH*, *PCCA*, and *DCLRE1C*. A large deletion of exon 18 has been identified in 5–8% of *GAA* alleles in patients with both infantile and adult-onset Pompe disease from diverse ethnic backgrounds.<sup>24</sup> This mutation occurs at even higher frequency in the Dutch population and has been proposed to be a founder mutation in this population.<sup>25</sup> Large deletions in the *PAH* gene, including the

deletion of exon 5 reported here (PAH consortium database; <http://www.pahdb.mcgill.ca>), have been found in ~0.5–0.8% of phenylketonuria patients.<sup>26</sup> In the *PCCA* gene, large deletions of one or more exons have been reported in association with propionic acidemia.<sup>27</sup> In one study of 66 patients with propionic acidemia, exonic deletions were identified on 21.2% of alleles of patients with no mutation identified by sequencing on one or both alleles.<sup>28</sup> For the *DCLRE1C* gene, in two studies the most frequent mutations were gross deletions of exons 1–3 or exons 1–4 due to a homologous recombination between the functional gene copy and a pseudogene located 61.2 kb upstream. Such deletions in *DCLRE1C* were reported in 59% of patients with mutations in one study and in all six of patients with mutations identified by sequencing in a second study of Saudi Arabian patients.<sup>29,30</sup>

Unlike in recessive disorders, intragenic copy number mutations in dominant disorders are more frequently recognized because a single mutation is sufficient to cause the phenotype. Translocations and intrachromosomal deletions have helped identify many causative genes in dominant disorders, e.g., *PAX6* in WAGR syndrome.<sup>31</sup> Deletions are common in some disorders, such as aniridia, von Hippel–Lindau syndrome or Peutz–Jeghers syndrome. We found a large number of copy number mutations in dominant disorders in our study, consistent with previous reports. This underscores the utility of exon-level array CGH for a broad range of disorders caused by haploinsufficiency (Table 1), including hereditary cancer predisposition syndromes such as Von Hippel–Lindau syndrome, Peutz–Jeghers syndrome, multiple hereditary exostoses, and Gorlin syndrome.

For some disorders, the frequency of copy number mutations in our data deviated from previously reported accounts. One explanation is that patients referred for testing do not always meet stringent clinical diagnostic criteria for the suspected disorder and may have another genetic disorder (phenotypic heterogeneity) or have mutations in another gene (genetic heterogeneity). In addition, fluorescence in situ hybridization may have been used to exclude large deletions in disorders in which such deletions are relatively frequent (e.g., disorders associated with *VHL* and *PAX6/WT1* mutations). Finally, because the number of patients studied for some of these disorders in our study or elsewhere was small, the observed frequency of copy number mutations in those genes may not be a true representation of the actual deletion/duplication rate.

An example of a discrepancy between published mutation rates and our data is deletions/duplications involving *PTEN*. Mutations in this gene are associated with multiple disorders, including CS, BRRS, PROTEUS syndrome, PROTEUS-like syndrome, and macrocephaly and autism spectrum disorders.<sup>32,33</sup> The published mutation rates range from 0.26% of all cases to 13% of sequencing-negative cases.<sup>34,35</sup> In our large cohort of 1,231 individuals, exonic copy number mutations of *PTEN* were found in 0.5% of individuals tested. Individuals with a whole or partial gene deletion or a partial gene duplication had CS, BRRS, and general overgrowth phenotypes, and none had a diagnosis



of macrocephaly or autism spectrum disorder. One previous study identified one whole and two partial *PTEN* mutations by MLPA in a carefully selected cohort of 122 CS or BRRS patients who were negative by gene sequencing, reflecting an overall deletion rate of 2.5%.<sup>8</sup> However, all patients with a *PTEN* deletion had a BRRS phenotype (3/27 or 11%). Another study examined 30 sequencing-negative patients with CS and found four deletion mutations involving exon 1 (13%).<sup>35</sup> By contrast, a large multicenter prospective study of 3,042 individuals with relaxed CS clinical criteria revealed large exonic deletions of *PTEN* in only 8 (0.26%).<sup>34</sup> These deletions accounted for <3% of all types of mutations identified. Results of this large study concur with our data (0.5% frequency) in an unselected cohort referred for *PTEN* testing. Our results support the current understanding that *PTEN* deletions are more common in BRRS and perhaps CS, but are not a general cause for autism spectrum disorders.<sup>36,37</sup>

We found several copy number mutations in X-linked genes that illustrate two advantages of using exon array CGH. First, it readily identified female carriers of a whole or partial gene deletion on the X chromosome. In contrast to males, in whom hemizygous deletions can be readily recognized by failure of PCR amplification, heterozygous gene deletions in females are not detectable by Sanger sequencing. Deletions are quite frequent in several X-linked disorders, including ectodermal dysplasia (~10%), OTC deficiency (~8–15%), and Rett syndrome (~10%).<sup>11</sup> The second advantage of the exon array is that it can also detect pathogenic intragenic duplications, in both males and females, that might otherwise go undetected by sequencing. The X chromosome contains many disease genes in which intragenic duplications have been reported, including *MECP2* (Rett syndrome), *RPS6KA3* (Coffin–Lowry syndrome), and *NR0B1* (46,XY gonadal dysgenesis).<sup>11</sup> We tested 7 males for X-linked disorders and found no duplications, but 1 out of 211 females was positive for a partial gene duplication in *MECP2*.

Our findings illustrate the utility of a robust exon array CGH method for detecting intragenic copy number mutations in Mendelian disorders. The data presented here emphasize the enhanced testing sensitivity for Mendelian disorders concurrently or after sequencing. They suggest that exon array CGH should be routinely used for autosomal recessive disorders, particularly IEMs that have only one mutation identified by sequencing or for conditions in which loss-of-function mutations or abnormal gene dosage explain the associated phenotype. Although we have targeted only 589 genes of known clinical significance and examined only a single gene or a small panel of genes in each individual, exon-level array CGH could also be designed as a screening test. For example, exon-level coverage of a few thousand genes may be appropriate for simultaneous analysis on a whole-genome cytogenetic array CGH design because many genes linked to developmental disorders arise from haploinsufficiency. It is also now feasible and relatively inexpensive to perform whole-exome copy number analysis by array CGH to complement exome sequencing until robust algorithms to calculate copy number from next-generation sequencing data become available.

## SUPPLEMENTARY MATERIAL

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

## DISCLOSURE

The authors are current or past employees of GeneDx.

## REFERENCES

- den Dunnen JT, White SJ. MLPA and MAPH: sensitive detection of deletions and duplications. *Curr Protoc Hum Genet* 2006;Chapter 7:Unit 7.14.
- Konialis C, Haghefelt B, Sevastidou S, et al. Uncovering recurrent microdeletion syndromes and subtelomeric deletions/duplications through non-selective application of a MLPA-based extended prenatal panel in routine prenatal diagnosis. *Prenat Diagn* 2011;31:571–577.
- Ravnan JB, Tepperberg JH, Papenhausen P, et al. Subtelomere FISH analysis of 11 688 cases: an evaluation of the frequency and pattern of subtelomere rearrangements in individuals with developmental disabilities. *J Med Genet* 2006;43:478–489.
- Miller DT, Adam MP, Aradhya S, et al. Consensus statement: chromosomal microarray is a first-tier clinical diagnostic test for individuals with developmental disabilities or congenital anomalies. *Am J Hum Genet* 2010;86:749–764.
- Tayeh MK, Chin EL, Miller VR, Bean LJ, Coffee B, Hegde M. Targeted comparative genomic hybridization array for the detection of single- and multiexon gene deletions and duplications. *Genet Med* 2009;11:232–240.
- Boone PM, Bacino CA, Shaw CA, et al. Detection of clinically relevant exonic copy-number changes by array CGH. *Hum Mutat* 2010;31:1326–1342.
- Orloff MS, Eng C. Genetic and phenotypic heterogeneity in the *PTEN* hamartoma tumour syndrome. *Oncogene* 2008;27:5387–5397.
- Zhou XP, Waite KA, Pilarski R, et al. Germline *PTEN* promoter mutations and deletions in Cowden/Bannayan-Riley-Ruvalcaba syndrome result in aberrant *PTEN* protein and dysregulation of the phosphoinositol-3-kinase/Akt pathway. *Am J Hum Genet* 2003;73:404–411.
- Hearle NC, Rudd MF, Lim W, et al. Exonic *STK11* deletions are not a rare cause of Peutz-Jeghers syndrome. *J Med Genet* 2006;43:e15.
- Jennes I, Entius MM, Van Hul E, Parra A, Sangiorgi L, Wuyts W. Mutation screening of *EXT1* and *EXT2* by denaturing high-performance liquid chromatography, direct sequencing analysis, fluorescence in situ hybridization, and a new multiplex ligation-dependent probe amplification probe set in patients with multiple osteochondromas. *J Mol Diagn* 2008;10:85–92.
- Pagon RA. GeneTests: an online genetic information resource for health care providers. *J Med Libr Assoc* 2006;94:343–348.
- Soufir N, Gerard B, Portela M, et al. *PTCH* mutations and deletions in patients with typical nevoid basal cell carcinoma syndrome and in patients with a suspected genetic predisposition to basal cell carcinoma: a French study. *Br J Cancer* 2006;95:548–553.
- Hoebeek J, van der Lijft R, Poppe B, et al. Rapid detection of *VHL* exon deletions using real-time quantitative PCR. *Lab Invest* 2005;85:24–33.
- Warthen DM, Moore EC, Kamath BM, et al. Jagged1 (*JAG1*) mutations in Alagille syndrome: increasing the mutation detection rate. *Hum Mutat* 2006;27:436–443.
- Kalmár L, Bors A, Farkas H, et al. Mutation screening of the C1 inhibitor gene among Hungarian patients with hereditary angioedema. *Hum Mutat* 2003;22:498.
- Chang EH, Menezes M, Meyer NC, et al. Branchio-oto-renal syndrome: the mutation spectrum in *EYA1* and its phenotypic consequences. *Hum Mutat* 2004;23:582–589.
- Robinson DO, Howarth RJ, Williamson KA, van Heyningen V, Beal SJ, Crolla JA. Genetic analysis of chromosome 11p13 and the *PAX6* gene in a series of 125 cases referred with aniridia. *Am J Med Genet A* 2008;146A:558–569.
- Redeker EJ, de Visser AS, Bergen AA, Mannens MM. Multiplex ligation-dependent probe amplification (MLPA) enhances the molecular diagnosis of aniridia and related disorders. *Mol Vis* 2008;14:836–840.
- Hardwick SA, Reuter K, Williamson SL, et al. Delineation of large deletions of the *MECP2* gene in Rett syndrome patients, including a familial case with a male proband. *Eur J Hum Genet* 2007;15:1218–1229.

20. Parri V, Katzaki E, Uliana V, et al. High frequency of COH1 intragenic deletions and duplications detected by MLPA in patients with Cohen syndrome. *Eur J Hum Genet* 2010;18:1133–1140.
21. Zhang F, Gu W, Hurler ME, Lupski JR. Copy number variation in human health, disease, and evolution. *Annu Rev Genomics Hum Genet* 2009;10:451–481.
22. Hastings PJ, Lupski JR, Rosenberg SM, Ira G. Mechanisms of change in gene copy number. *Nat Rev Genet* 2009;10:551–564.
23. Weaver DD, Solomon BD, Akin-Samson K, Kelley RI, Muenke M. Cyclopia (synophthalmia) in Smith-Lemli-Opitz syndrome: First reported case and consideration of mechanism. *Am J Med Genet C Semin Med Genet* 2010;154C:142–145.
24. Joshi PR, Gläser D, Schmidt S, et al. Molecular diagnosis of German patients with late-onset glycogen storage disease type II. *J Inherit Metab Dis* 2008;31 Suppl 2:S261–S265.
25. Kroos MA, Van der Kraan M, Van Diggelen OP, et al. Glycogen storage disease type II: frequency of three common mutant alleles and their associated clinical phenotypes studied in 121 patients. *J Med Genet* 1995;32:836–837.
26. Gable M, Williams M, Stephenson A, et al. Comparative multiplex dosage analysis detects whole exon deletions at the phenylalanine hydroxylase locus. *Hum Mutat* 2003;21:379–386.
27. Yang X, Sakamoto O, Matsubara Y, et al. Mutation spectrum of the PCCA and PCCB genes in Japanese patients with propionic acidemia. *Mol Genet Metab* 2004;81:335–342.
28. Desviat LR, Sanchez-Alcudia R, Pérez B, et al. High frequency of large genomic deletions in the PCCA gene causing propionic acidemia. *Mol Genet Metab* 2009;96:171–176.
29. Pannicke U, Hönig M, Schulze I, et al. The most frequent DCLRE1C (ARTEMIS) mutations are based on homologous recombination events. *Hum Mutat* 2010;31:197–207.
30. Alsmadi O, Al-Ghonaïm A, Al-Muhsen S, et al. Molecular analysis of T-B-NK+ severe combined immunodeficiency and Omenn syndrome cases in Saudi Arabia. *BMC Med Genet* 2009;10:116.
31. Dallapiccola B, Mingarelli R, Novelli G. The link between cytogenetics and mendelism. *Biomed Pharmacother* 1995;49:83–93.
32. Eng C. PTEN: one gene, many syndromes. *Hum Mutat* 2003;22:183–198.
33. Varga EA, Pastore M, Prior T, Herman GE, McBride KL. The prevalence of PTEN mutations in a clinical pediatric cohort with autism spectrum disorders, developmental delay, and macrocephaly. *Genet Med* 2009;11:111–117.
34. Tan MH, Mester J, Peterson C, et al. A clinical scoring system for selection of patients for PTEN mutation testing is proposed on the basis of a prospective study of 3042 probands. *Am J Hum Genet* 2011;88:42–56.
35. Pezzolesi MG, Zbuk KM, Waite KA, Eng C. Comparative genomic and functional analyses reveal a novel cis-acting PTEN regulatory element as a highly conserved functional E-box motif deleted in Cowden syndrome. *Hum Mol Genet* 2007;16:1058–1071.
36. Orrico A, Galli L, Buoni S, Orsi A, Vonella G, Sorrentino V. Novel PTEN mutations in neurodevelopmental disorders and macrocephaly. *Clin Genet* 2009;75:195–198.
37. Buxbaum JD, Cai G, Chaste P, et al. Mutation screening of the PTEN gene in patients with autism spectrum disorders and macrocephaly. *Am J Med Genet B Neuropsychiatr Genet* 2007;144B:484–491.
38. The Retinoschisis Consortium. Functional implications of the spectrum of mutations found in 234 cases with X-linked juvenile retinoschisis. *Hum Mol Genet* 1998;7:1185–1192.