

Enhanced detection of clinically relevant genomic imbalances using a targeted plus whole genome oligonucleotide microarray

Erin L. Baldwin, PhD, Ji-Yun Lee, PhD, Douglas M. Blake, BS, Brian P. Bunke, BS, Chad R. Alexander, BS, Amy L. Kogan, BS, David H. Ledbetter, PhD, and Christa L. Martin, PhD

Purpose: Array comparative genomic hybridization is rapidly becoming an integral part of cytogenetic diagnostics. We report the design, validation, and clinical utility of an oligonucleotide array which combines genome-wide coverage with targeted enhancement at known clinically relevant regions. **Methods:** Probes were placed every 75 kb across the entire euchromatic genome to establish a chromosomal “backbone” with a resolution of ~500 kb, which is increased to ~50 kb in targeted regions. **Results:** For validation, 30 samples showed 100% concordance with previous G-banding and/or fluorescence in situ hybridization results. Prospective array analysis of 211 clinical samples identified 33 (15.6%) cases with clinically significant abnormalities. Of these, 23 (10.9%) were detected by the “targeted” coverage and 10 (4.7%) by the genome-wide coverage (average size of 3.7 Mb). All abnormalities were verified by fluorescence in situ hybridization, using commercially available or homebrew probes using the 32K bacterial artificial chromosome set. Four (1.9%) cases had previously reported imbalances of uncertain clinical significance. Five (2.4%) cases required parental studies for interpretation and all were benign familial variants. **Conclusions:** Our results highlight the enhanced diagnostic utility of a genome-wide plus targeted array design, as the use of only a targeted array would have failed to detect 4.7% of the clinically relevant imbalances. **Genet Med 2008;10(6):415–429.**

Key Words: array comparative genomic hybridization, oligonucleotide microarray, copy number variant, molecular karyotype, genome-wide

The G-banded karyotype was the first genome-wide analysis tool and has been the “Gold Standard” for detection of aneuploidy and segmental aneuploidy for both syndromic and idiopathic mental retardation (MR) since its introduction in the early 1970s. The standard G-banded karyotype at a 550 band resolution¹ has generally been thought to accurately detect deletions and duplications of ≥ 5 Mb with a yield of 5–10% abnormalities, depending on ascertainment criteria.^{2,3} The introduction of high-resolution chromosome analysis⁴ made it possible to discover new “microdeletion syndromes” caused by smaller deletions (3–5 Mb). Although some laboratories implemented high-resolution techniques for genome-wide analysis in idiopathic MR, its primary application has been in

“focused” or targeted analysis of specific chromosomal regions suspected due to clinical features consistent with a specific microdeletion syndrome. Further development of high-resolution chromosome banding demonstrated the potential for a 2000 band resolution⁵; however, this method has not been practical in a diagnostic laboratory setting.

Targeted application of high-resolution banding for microdeletion syndromes was replaced in the early 1990s with the advent of locus-specific fluorescence in situ hybridization (FISH) analysis using large genomic clones (cosmids or BACs), e.g., for Miller-Dieker syndrome⁶ (MDLS [MIM 247200]) (OMIM, <http://www.ncbi.nlm.nih.gov/sites/entrez?db=OMIM>), Prader-Willi/Angelman syndromes⁷ (PWS/AS [MIM 176270/105830]) and 22q11 deletion syndrome⁸ (DGS [MIM 188400]). The development of a complete set of unique FISH clones near the junction of subtelomeric sequences of each chromosome arm^{9,10} allowed a targeted analysis of all human telomeres to identify cryptic deletions, duplications, and unbalanced translocations. Application of these probes in patients with idiopathic MR and normal G-banded karyotypes identifies an additional 3–6% of abnormalities,^{10–12} and “telomere FISH” quickly became a standard-of-care test following, or in conjunction with, G-banded chromosome analysis.^{2,3}

Comparative genomic hybridization (CGH) on metaphase chromosomes was initially introduced as a genome-wide method to detect copy number gains and losses of ≥ 10 Mb in

From the Department of Human Genetics, Emory University, Atlanta, Georgia.

Christa Lese Martin, PhD, FACMG, Emory University School of Medicine, 615 Michael St., Suite 301, Atlanta, GA 30322. E-mail: clmartin@genetics.emory.edu.

The first two authors contributed equally to this work.

Disclosure: The authors declare no conflict of interest.

E.L.B. and J.-Y.L. contributed equally to this manuscript.

Supplementary tables are available via the ArticlePlus feature at www.geneticsinmedicine.org. Please go to the June issue and click on the ArticlePlus link posted with the article in the Table of Contents to view this material.

Submitted for publication October 31, 2007.

Accepted for publication February 27, 2008.

DOI: 10.1097/GIM.0b013e318177015c

solid tumors.¹³ At this low resolution, CGH was not useful for research or clinical application in pediatric or prenatal cytogenetic investigations where G-banding analysis provided better resolution. However, the change in the hybridization target for CGH from normal metaphase chromosomes to DNA from genomic clones allowed the detection of single copy changes (deletion or duplication) at the level of individual BAC clones (~100 kb).^{14,15}

Targeted BAC arrays, comprised of several hundred to a thousand or more clones representing human telomeres and microdeletion/microduplication syndromes, have proven to be a reliable and sensitive method for identifying genomic imbalances in idiopathic MR as an adjunct to conventional G-banded karyotype, with a yield of approximately 5–10%.^{16–18} An obvious drawback to such targeted arrays are the large gaps in genome coverage which, by design, will not detect large-scale imbalances (~5 Mb or more) for the many genomic regions that do not include targeted clinically relevant loci.

One approach to genome-wide coverage has been to place BAC clones every 1 Mb throughout the genome,^{19,20} and some clinical cytogenetics laboratories have used this genome-wide array strategy in combination with a targeted BAC array to capitalize on the advantages of these two approaches. Such a 1 Mb array is approximately equivalent to a 3000 band karyotype, although the reliability and clinical significance of a single BAC clone gain or loss are uncertain without further confirmation and follow-up using other methods.

Development and application of high-resolution BAC arrays containing over 32,000 tiling-path (overlapping) BAC clones have been reported by several groups^{21–23} and applied in cancer and in children with unexplained MR. Although these tiling path BAC arrays have a potential resolution of ~50 kb,²⁴ one group²³ used a threshold of at least three adjacent clones (~300–400 kb) to indicate potential clinical significance. These authors argue that the benefit of detecting an additional ~5% of clinically significant abnormalities by such a whole genome approach outweighs the additional work and potential anxiety of identifying a small percentage of benign copy number variations (CNVs),^{25,26} which are very common in the human genome.^{27–34}

An alternative to genomic or polymerase chain reaction-amplified DNA from large-insert clones has been the use of synthetic oligonucleotides of 25–85 mer length for genome-wide copy number assessment. Recent reviews of the various platforms available and their performance characteristics are available.^{35,36} Several manufacturers have developed genome-wide array comparative genomic hybridization (aCGH) microarrays containing ~40,000–400,000 probes covering the human genome.^{37,38} This technology has been used in the detection of chromosomal imbalances associated with MR or autism.^{39,40} However, these commercial arrays were not designed specifically for clinical cytogenetics applications, and therefore may not have adequate density and coverage for clinically relevant regions, such as telomeres, centromeres and microdeletion or microduplication syndromes.^{41,42} To overcome this deficiency, it is possible to design custom microarrays to enhance

the coverage at clinically relevant regions. Recently, Toruner et al.⁴¹ modified a commercial microarray by removing one-third of the oligonucleotides in the backbone regions and adding enriched probe coverage for the telomere regions.

Genome-wide aCGH can thus be accomplished with high-resolution BAC or oligonucleotide probe coverage, and either will likely identify an additional 3–5% of clinically significant genomic imbalances as compared with a targeted array approach. However, concerns have been expressed regarding the readiness for clinical application because of the potentially high rate of copy number variants of unknown clinical significance.^{43,44}

It would be ideal if one could increase the total yield of clinically significant imbalances but minimize the frequency of follow-up studies to determine the clinical significance of imbalances in regions of the genome. One approach to this optimization strategy is to use a size threshold as an initial screen to differentiate those genomic copy number differences most likely to have clinical significance from those most likely to be benign variants. This single criterion is obviously not absolute, but can be combined with additional information including family studies (to differentiate *de novo* from inherited events), gene content, and data from the CNV databases and published literature to allow a strong clinical prediction in most cases.

Here, we describe an oligonucleotide-based aCGH design which combines the strengths of a targeted array containing clinically relevant loci with the enhanced detection potential of a genome-wide array. The validation and initial prospective application of this array are presented and demonstrate that the use of such a “molecular karyotype,” equivalent to a 6000 band karyotype, significantly enhances the detection of clinically significant deletions and duplications not detected by G-banding or current targeted array designs.

MATERIALS AND METHODS

Patient samples

A protocol approved by the Institutional Review Board at Emory University was followed for this study and informed consent was obtained as required by the institutional review board. A retrospective analysis of 30 patient samples was performed for validation (Supplementary Table 1). An initial validation study was performed on 10 patients with known chromosomal abnormalities. Subsequently, a blinded analysis of 20 patients was performed, including individuals with normal karyotypes and various chromosome abnormalities previously identified by G-banding, locus-specific FISH and/or subtelomeric FISH.

Following this validation, prospective studies were performed on clinical samples referred for various indications including unexplained developmental delay, MR, dysmorphic features, congenital anomalies, autism or clinical features suggestive of a chromosomal syndrome.

DNA was isolated from peripheral blood samples collected in ethylenediaminetetraacetate tubes using a commercially available DNA extraction kit (Puregene blood kit, QIAGEN Inc., Valencia, CA). The manufacturer’s recommended proto-

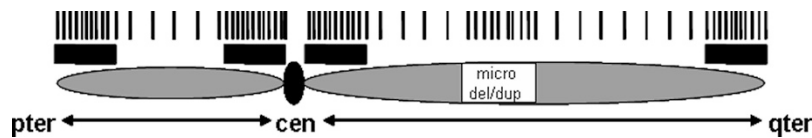


Fig. 1. Schematic of custom array design. The general design of the custom array is shown where the vertical lines represent the oligonucleotides: one probe is placed every ~75 kb for the backbone coverage. Other clinically relevant regions, including telomeres, centromeres, common microdeletion/microduplication syndromes and selected Mendelian disorders, contain enhanced coverage. The black bars represent the unique telomere FISH clones and the unique centromere FISH clones that were also targeted on the array.

col was followed without modification. For array confirmation studies (FISH or G-banding), peripheral blood cultures were established and harvested after 72-hours following a routine cytogenetic protocol.

Array design

To develop a high-resolution molecular karyotype, we designed an oligonucleotide microarray that combines targeted with genome-wide coverage (Fig. 1). This array contains 43,102 oligonucleotide probes in a $4 \times 44K$ format (Agilent Technologies, Santa Clara, CA). Using the eArray v4.0 tool (Agilent), we selected 42,869 probes specific for the human genome (UCSC hg17 May 2004; NCBI Build 35) to create both targeted and genome-wide coverage on the array. Probes were selected by searching genomic intervals using eArray's High Density Comparative Genomic Hybridization search tool, with a preference for probes that were located in unique sequences (avoiding segmental duplications) and were validated by the manufacturer.

Using eArray, an oligonucleotide backbone was established across the euchromatic regions of the genome, including 41,023 probes spaced on average every 75 kb. Using a sliding window of 4–5 oligonucleotides, this backbone coverage provides a resolution of ~300 kb. However, as discussed below (see Clinical Interpretation section), a threshold of 500 kb was used to minimize the detection of benign copy number variants and to enrich for clinically significant imbalances. This 500 kb resolution equates to a 6000 band karyotype capability (based on ~3000 Mb of DNA in the human genome).

Both pseudoautosomal regions (PAR1 and PAR2) were covered with a total of 233 oligonucleotides. Probes were selected to include coverage for the known genes in these regions; the remaining oligonucleotides were evenly distributed among PAR1 (chrX:1–2,702,389) and PAR2 (chrX:154,472,117–154,900,000) (NCBI Build 35).

For targeted coverage, 1846 probes were selected using eArray to create additional high density coverage of oligonucleotides in the known clinically relevant regions including the telomeres, centromeres, common microdeletion/microduplication syndromes and selected Mendelian disorders (Supplementary Table 2). The oligonucleotides selected to cover these targeted regions were overlaid onto the backbone coverage to create an array with targeted and genome-wide coverage. Any overlapping probes were removed and replaced with neighboring oligonucleotides to enhance the coverage in the region.

To enhance the coverage of the telomere regions, we used our previous physical mapping data, and bioinformatics re-

sources, to identify a BAC clone that defined the transition from the subtelomeric repeats to the unique DNA of each chromosome arm.^{9,10,42,45} Briefly, clones were selected based on physical maps⁴⁶ and publicly available resources, such as the UCSC Genome Browser (<http://genome.ucsc.edu>),^{47,48} the Whitehead Institute STS/radiation hybrid maps (http://www-genome.wi.mit.edu/cgi-bin/contig/phys_map), and the NCBI Map Viewer (<http://www.ncbi.nlm.nih.gov>). FISH mapping was used to determine the junction between the subtelomeric repeats and unique DNA. Clones located in the subtelomeric repeats display cross-hybridization and, therefore, were avoided in the design of the custom array. The most distal clones that showed no cross-hybridization were selected as the unique telomere clone. Using the location of these 41 unique telomere clones, a minimum of 10 oligonucleotides was used to represent each clone so that small imbalances could be detected.

To enhance the coverage of the centromere regions, we identified the BAC clones that defined the transition from the pericentromeric repeats to the unique DNA for each chromosome arm.⁴⁹ Briefly, fully sequenced clones that contained <10% duplicated sequence using the segmental duplication track⁵⁰ were initially chosen from NCBI Build 33 using the UCSC genome browser database, whereas subsequent clones were chosen based on the most current NCBI builds.^{47,48,51} FISH mapping was performed to identify the most proximal clone to each pericentromeric repeat-unique DNA junction for each chromosome arm. Clones located within the pericentromeric repeats display cross-hybridization, and these regions were avoided in the design of the array. The most proximal clone for each centromere arm was confirmed by FISH to map uniquely and to the correct position. Using the location of these 43 unique centromere FISH clones, a minimum of 10 oligonucleotides for each clone was added to the array.

For each specific clinical region of interest, including the known microdeletion/microduplication syndromes and selected Mendelian disorders (Supplementary Table 2), an average of 10 oligonucleotides was added to the backbone coverage. To maximize the number of oligonucleotides, probes were selected using both the genomic interval and gene annotation options in eArray. If the probes were located within a segmental duplication, adjacent oligonucleotides were chosen to avoid these duplicated or repetitive regions.

Array CGH

aCGH was performed following the manufacturer's protocol (Agilent Technologies). As an internal hybridization control for each experiment, patient samples were matched with a

pooled normal control DNA sample of the opposite sex (Promega Corporation, Madison, WI). Genomic DNA (0.5–1.5 μg) was digested with *AluI* and *RsaI* (Promega Corporation, Madison, WI) for 2 hours. The DNA was labeled for 2 hours using random primers, Cy-3 and Cy-5-dUTP dyes and Exo-Klenow fragment (Agilent Technologies). Patient DNA (labeled with Cy-3) was combined with normal control DNA (labeled with Cy-5) and hybridized to the array in the presence of Cot-1 DNA (Invitrogen Corporation, Carlsbad, CA). After a 24-hour hybridization at 65°C, the slides were washed using the “Wash Procedure B with Stabilization and Drying Solution” protocol and scanned using a GenePix Autoloader 4200AL or GenePix 4000B (Molecular Devices, Sunnyvale, CA).

Data analysis

Two different software packages were used for data analysis. BlueFuse software (BlueGnome Ltd., Cambridge, UK) was initially used during our array validation studies. A second software package, Feature Extraction 9.5.11 and CGH Analytics 3.4.40 (Agilent Technologies), was later used to re-evaluate the validation study data and for prospective analyses. In a comparison of the array results from BlueFuse and CGH Analytics, no discrepancies were observed in the ability to detect chromosomal imbalances using the settings described.

For BlueFuse, normalization of the data were performed using Block Lowess, which corrects for intensity-related variation within images. Regions of copy number alterations were detected using set thresholds for the channel ratios based on three standard deviations from the median of all autosomes. The thresholds for the \log_2 ratios were set at -0.32 for losses and 0.26 for gains. Channel 1 (Ch1) represented the patient sample and channel 2 (Ch2) represented the normal control DNA.

The software was set to flag a potential abnormality when the minimum number of contiguous probes showing loss or gain of intensity was 4–5 oligonucleotides. With a backbone coverage of probes spaced every 75 kb, this setting allows the detection of abnormalities of ~ 300 kb. For the targeted regions, the minimum of five oligonucleotides allows for the detection of ~ 50 kb imbalances. The size of the imbalance (as indicated in Tables 1 and 2) includes the most distal and most proximal oligonucleotides that are contained within the loss/gain.

For data analysis using Feature Extraction 9.5.11 and CGH Analytics 3.4.40, the ADM-1 statistical algorithm in CGH Analytics was used with a sensitivity threshold of 6.7. In order for CGH Analytics to call an abnormality, the minimum number of probes included in a region of deletion or amplification was set at four oligonucleotides. The thresholds for the \log_2 ratios were set at -0.32 for losses and 0.26 for gains.

Gene content

The UCSC genome browser (May 2004) was used to assess the genomic architecture of the abnormal region (Segmental Duplications track)⁵⁰ and to assess the number of known genes (UCSC Known genes track) contained within the imbalance.^{52,53} The “Segmental Dups” track displays regions of DNA that are >1 kb with at least 90% sequence homology with

another region. The genes were displayed in four colors to represent the level of supporting data: Black, entry in the Protein Databank (PDB); Dark Blue, either a corresponding RefSeq mRNA that is “Reviewed” or “Validated” or a corresponding Swiss-Prot protein; Medium Blue, corresponding RefSeq mRNA that is not “Reviewed” nor “Validated”; Light Blue, no corresponding PDB entry, RefSeq mRNA, or Swiss-Prot protein. For this study, the number of known genes included the genes that have an entry in the PDB (Black) or a validated RefSeq mRNA or Swiss-Prot protein entry (Dark Blue).

FISH confirmation studies

If a chromosomal abnormality was detected by aCGH, follow-up confirmation studies were performed. The majority of abnormalities were confirmed by FISH studies (a few cases were confirmed by G-banding when array and G-banding analyses were completed simultaneously). For abnormalities detected in the telomeres or the common microdeletion/microduplication regions, commercially available FISH probes (Abbott Molecular Inc., Des Plaines, IL) were used.

For all other regions in which a commercial probe was not available, homebrew FISH probes were custom designed specific to the region in which the imbalance was detected. Clones were selected from the “Human genome high-resolution BAC rearranged clone set; 32K set,” which was originally developed and described by Krzywinski et al.²² and is available from the BACPAC Resource Center (Children’s Hospital Oakland Research Institute, Oakland, CA). This clone set consists of BAC clones that were ordered into tiling path coverage of the human genome based on restriction digest fingerprinting.

The complete clone set is maintained locally to provide immediate access as needed. The coordinates of the oligonucleotides that show either a gain or a loss on the array are matched to a corresponding BAC clone from the 32K set. Clone DNA is labeled using a random priming reaction. Briefly, clone DNA and random primers (Invitrogen Corporation) were denatured for 7 minutes. On ice, dNTPs, fluorescently-labeled nucleotides [Spectrum Orange-dUTP, Spectrum Green-dUTP] (Abbott Molecular Inc., Des Plaines, IL), and Klenow (Invitrogen Corporation) were incorporated and incubated at 37°C for 2 hours. Probes were precipitated and subsequently used in standard FISH assays.^{45,54} All probes were tested on normal controls in addition to the patient samples.

Evaluation of common CNVs

We analyzed the size distribution of CNVs contained in the Database of Genomic Variants (<http://projects.tcag.ca/variation>)²⁷ as of October 2006 by downloading all CNV data into Excel spreadsheets and sorting by size in 100 kb bins. Approximately 50% of all CNVs are smaller than 100 kb, and $\sim 95\%$ are <500 kb. In addition, a recent large study using a 500K SNP array, reported a median CNV size of 81 kb.³² Given these data, we chose to use a threshold of 500 kb for the detection size of an imbalance located *outside of the known clinically relevant regions* to enrich for clinically significant imbalances and decrease the number of benign CNVs detected.

Table 1
Abnormal results detected by the targeted coverage

Case	aCGH ISCN nomenclature (NCBI build 35)	Size loss (Mb)	Size gain (Mb)	Clinical target
2621 ^a	arr cgh 11q24.2q25(126,979,765 bp → 134,373,558 bp) × 1	7.4		Telomere
2752	arr cgh 15q11.2q13.1(21,208,377 bp → 26,232,997 bp) × 1	5.0		PWS/AS
2911	arr cgh Xq28(152,715,644 bp → 153,116,375 bp) × 3		0.4	Rett syndrome
3129	arr cgh 9q34.3(137,948,336 bp → 137,988,441 bp) × 1	0.04		<i>EHMT1</i>
3202	arr cgh 21q22.3qter(42,526,105 bp → 46,914,745 bp) × 1	4.4		Telomere
3263	arr cgh 7q36.2qter (154,236,380 bp → 158,427,693 bp) × 1	4.2		Telomere
3279	arr cgh 1pterp36.32(819,590 bp → 3,259,116 bp) × 1	2.4		Telomere
3360 ^a	arr cgh 11q23.3qter(116,227,321 bp → 134,373,617 bp) × 3, 22q11.1q11.21 (15,772,052 bp → 18,686,317 bp) × 3		18.1 2.9	Telomere 22q11.2
3416 ^a	arr cgh 11q24.1qter(123,531,640 bp → 134,373,617 bp) × 1, 12q24.23qter(117,065,984 bp → 132,369,258 bp) × 3	10.8	15.3	Telomere
3483	arr cgh 15q11.2q13.1(21,208,377 bp → 26,193,908 bp) × 1	5.0		PWS/AS
3498	arr cgh 22q11.2(17,036,278 bp → 19,886,068 bp) × 1	2.8		22q11.2
3595	arr cgh 15q11.2(23,133,513 bp → 23,178,075 bp) × 1mat	0.04		AS
3603 ^a	arr cgh 7q32.1qter(128,543,958 bp → 158,574,555 bp) × 3, 20pterp13(18,580 bp → 1,781,123 bp) × 1	1.8	30.0	Telomere
3611	arr cgh 3q29(197,228,667 bp → 198,805,413 bp) × 1	1.6		Telomere
3622	arr cgh 1pter1p36.32(819,590 bp → 3,651,868 bp) × 1, 19q13.33qter(55,386,288 bp → 63,738,997 bp) × 3	2.8	8.4	Telomere
3623	arr cgh Xq26.2q26.3(132,769,230 bp → 134,070,559 bp) × 1dn	1.3		Lesch-Nyhan
3725	arr cgh 5pterp15.33(75,149 bp → 3,457,358 bp) × 1, 5p15.33p15.32(3,556,104 bp → 4,633,604 bp) × 3	3.4	1.1	Telomere
3745	arr cgh 17p11.2(16,723,271 bp → 18,814,979 bp) × 3		2.1	17p11.2
3783 ^a	arr cgh Xp22.33q28(3,003,888 bp → 154,492,983 bp) × 3, Yp11.31q12(2,698,214 bp → 57,369,605 bp) × 1		Whole Chromosome	Sex chromosome Aneuploidy
3787	arr cgh 15q11.2q13.1(21,258,545 bp → 26,193,908 bp) × 1	4.9		PWS/AS
3864 ^a	arr cgh 13q14.12q31.3(44,439,658 bp → 92,188,422 bp) × 1	47.7		<i>RB</i>
3919	arr cgh 22q11.21(18,984,101 bp → 19,886,068 bp) × 1	0.9		22q11.2
3922	arr cgh 22q11.21(17,279,496 bp → 19,829,971 bp) × 1	2.6		22q11.2

^aG-banding and microarray analyses were performed simultaneously or only an array study was requested.

Table 2
Abnormal results detected by the genome-wide backbone coverage

Case	aCGH ISCN nomenclature (NCBI build 35)	Size loss (Mb)	Size gain (Mb)	# Known genes
2026	arr cgh 2q11.2q12.3(100,453,490 bp → 108,547,845 bp) × 3		8.1	25
2818	arr cgh 2p16.1p15(59,797,632 bp → 62,474,869 bp) × 1	2.7		8
3484	arr cgh 12q21.2q21.31(74,363,745 bp → 79,091,897 bp) × 1mat	4.7		12
3596	arr cgh 16p13.3p13.2(5,957,777 bp → 6,619,303 bp) × 3		0.7	1
3715	arr cgh 15q21.3q22.2(55,535,566 bp → 59,986,776 bp) × 1	4.5		21
3841	arr cgh 2q31.1q31.2(173,470,111 bp → 178,288,235 bp) × 1	4.8		28
3882	arr cgh 14q23.1q23.3(59,453,522 bp → 66,287,274 bp) × 1dn	6.8		33
3963	arr cgh 16p11.2(29,500,284 bp → 30,098,069 bp) × 1	0.6		24
4153	arr cgh 9q33.3-q34.11(125,923,428 bp → 127,623,518 bp) × 1	1.7		12
4154	arr cgh 15q13.2q13.3(28,441,169 bp → 30,686,850 bp) × 1	2.2		6

Clinical interpretation

After array data review and confirmation of abnormalities by FISH or G-banding, the clinical significance of each imbalance was determined using criteria similar to that of routine G-banded cytogenetic analysis. A similar approach was also outlined in recently published guidelines for clinical molecular karyotyping studies.⁵⁵

Imbalances involving a known clinically relevant microdeletion/duplication region were considered abnormal and parental samples were requested only to rule out that one of the parents carried a balanced form of the rearrangement (with the exception of cases where one of the parents also showed an abnormal phenotype [e.g., 22q11 deletion]).

Imbalances that involved large (>2–3 Mb) genomic segments from the chromosomal backbone coverage, were considered to be likely pathogenic if they did not overlap a confirmed CNV region and contained multiple known genes. For these cases, parental analysis was recommended to rule out the possibility that one of the parents carried a balanced form of the rearrangement.

Imbalances that were <2 Mb in size were first evaluated to determine whether they overlapped a known CNV region. Imbalances were characterized as known CNVs if the loss or gain was included in one of the normal structural variation databases (Database of Genomic Variants²⁷ and the UCSC genome browser^{47,48}) as a copy number variant region with a population frequency of $\geq 1\%$ observed in multiple studies. Parental samples were not requested for these cases, because these are confirmed, common structural variations.

If an imbalance was <2 Mb in size and was in a unique genomic region, parental samples were requested to determine whether the imbalance was an inherited or a de novo event and the gene content was examined to determine the potential clinical significance of the finding.

RESULTS

Validation studies

We developed an oligonucleotide array that combines targeted with genome-wide coverage to produce a 6000 band “molecular karyotype” (Fig. 1). For the initial validation of this array, we tested 10 patients with known chromosomal abnormalities. The abnormalities included two supernumerary marker chromosomes, five telomere deletions, two unbalanced translocations, and a microdeletion involving 15q11–13 (Supplementary Table 1, Cases 1–10). For all 10 cases, the array accurately detected the known chromosomal abnormality. In addition to elucidating the precise size of the imbalance, the array results also revealed additional complex rearrangements in two cases (Cases 3 and 5, data not shown).

Interestingly, one case (Case 6) had a 1 Mb deletion of the 4q telomere region that was originally detected by telomere FISH analysis. This case had previously been analyzed using a genome-wide off-the-shelf commercial oligonucleotide array which failed to detect the imbalance because of poor genomic

coverage at this telomere region.⁴² Using the custom designed array, with enhanced coverage of the telomere regions, this 1 Mb deletion was accurately detected.

For further validation, a blinded analysis was performed on 20 samples, including both abnormal and normal cytogenetic findings (Supplementary Table 1, Cases 11–30). The chromosomal abnormalities included two cases with 22q11 deletion, one case with 15q11-q13 deletion, two duplications (involving 10q26 and 17p11.2), three subtelomeric deletions (9qter, 12qter, and 21qter), two cytogenetically visible unbalanced translocations [der (6) t (6;10) (q25.3;q24.33) and der (18) t (15;18) (q11.2;p11.2)], one patient with trisomy 21, one patient with trisomy 21 and XXY, one patient with a subtelomeric deletion (1pter) and XXY, and one patient with XYY. For all 14 abnormal cases, the known chromosomal imbalance was accurately detected. The remaining six samples displayed normal aCGH results, which matched the normal results from previous G-banding and/or FISH analyses.

Overall, the aCGH validation results for all 30 samples, including the 10 initial cases and the 20 cases in the blinded analysis, were consistent with previous FISH and/or karyotype findings, demonstrating 100% concordance for the detection of copy number imbalances.

Prospective application of the targeted plus genome-wide array

Following validation of the array, we prospectively analyzed 211 samples submitted for clinical testing. Abnormalities of clinical significance were detected in 33 samples (15.6%) (Tables 1 and 2), whereas 169 samples showed normal results by aCGH (80.1%). Four (1.9%) cases had previously reported imbalances of uncertain clinical significance. Five (2.4%) cases required parental studies for interpretation, all of which were subsequently interpreted as benign familial variants. Each of these three categories of imbalance is presented in more detail in the following sections.

Clinically significant imbalances detected by targeted coverage

Of the 33 abnormal samples, 23 (10.9% of all cases) were identified by the targeted coverage of the custom array (Table 1 and Fig. 2). All imbalances were confirmed by FISH with commercially available probes. The most frequent class of imbalance involved the telomere regions, as 10 cases (4.7%) were detected by the targeted coverage of the telomere regions. The telomeric losses and gains ranged in size from 1.1 to 30 Mb. This large variation in size demonstrates the added benefit of including genome-wide coverage on the array, because the size of the imbalance can be accurately determined in a single assay without a preset limitation on coverage at the telomere regions.

An example of an unbalanced translocation involving two telomere regions is shown in Figure 2, A. The patient (Case 3416) presented with thrombocytopenia and multiple congenital anomalies. Microarray analysis revealed a loss of the 11q telomere region (involving bands q24.1 to qter) and a gain of the 12q telomere region (involving bands q24.23 to qter). The deletion of chromosome 11q was ~10.8 Mb in size and the duplication of chromosome 12q was ~15.3 Mb in size. Dele-

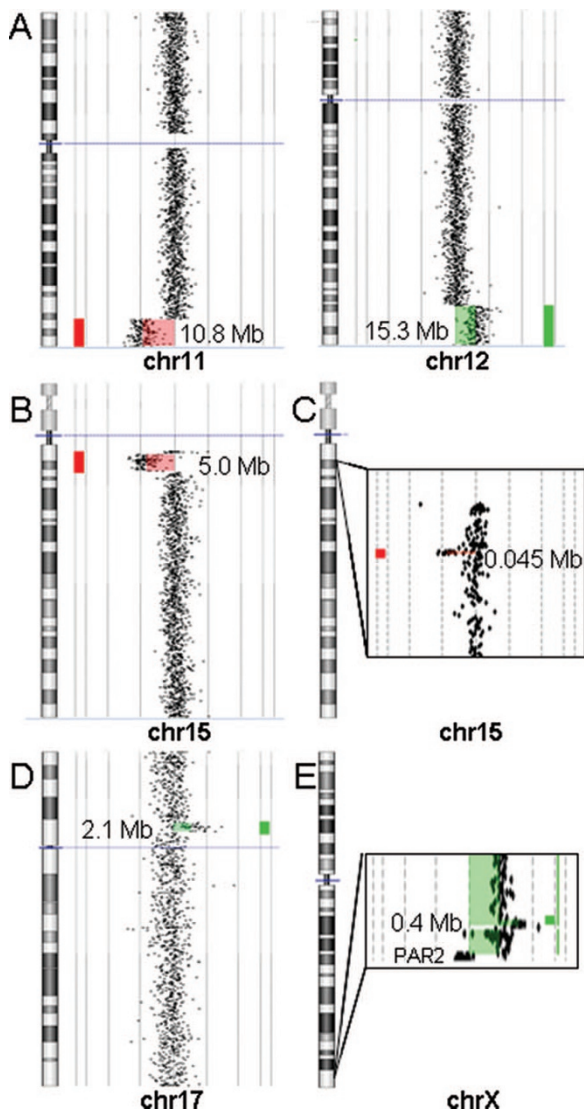


Fig. 2. Examples of imbalances detected by the targeted coverage of the custom microarray. The Y-axis displays the genomic position of the oligonucleotides for each chromosome, and the X-axis displays the \log_2 ratios of the patient sample versus a normal control sample. The center vertical line represents a \log_2 ratio of 0, with the neighboring lines indicating ratio values in increments of 1. Each dot represents a single oligonucleotide probe. A loss is depicted with a red bar, whereas a gain is depicted with a green bar. A, Case 3416 with an unbalanced telomeric translocation between the long arm of chromosomes 11 and 12 resulting in monosomy for a 10.8 Mb region on 11q and trisomy for a 15.3 Mb region on 12q; B, Case 2752 with a typical deletion of the PWS/AS critical region on 15q; C, Case 3595 with a 45 kb intragenic deletion of the *UBE3A* gene within the PWS/AS critical region; D, Case 3745 with a 2.1 Mb duplication involving the 17p11.2 region; E, Case 2911 with a 400 kb duplication of the *MECP2* gene on the long arm of the X chromosome.

tions of distal 11q cause Jacobsen syndrome (JBS [MIM 147791]), which includes thrombocytopenia as one of the cardinal phenotypic features. FISH analysis using probes for the 11q and 12q telomere regions demonstrated that these imbalances were the result of an unbalanced translocation [der (11) t (11;12) (q24.1;q24.23)] (data not shown). Although the monosomic and trisomic regions in this case were quite large, the net change in size of this rearrangement was ~ 4.5 Mb (which is at or below the resolution of routine G-banding anal-

ysis) and demonstrates how imbalances such as these can be easily missed by G-banding even though the individual segments involved in the translocation are quite large.

In addition to the telomeric imbalances, 14 abnormalities were detected with the targeted coverage of the other clinically relevant loci including the microdeletion/duplication syndromes and the Mendelian disorder regions (note that Case 3360 had two abnormalities: one at the 11q telomere region and one in the 22q11.2 targeted region). Two examples of deletions involving the PWS/AS critical region are shown in Figure 2. In Case 2752, a typical 5.0 Mb deletion of the PWS/AS region, including the *SNRPN* [MIM 182279] and *UBE3A* [MIM 601623] genes, was observed (Fig. 2, B). In addition to detecting the deletion, array analysis provided immediate breakpoint (BP) localization: the BPs in this case were shown to be at BP2 and BP3 of the PWS/AS region, consistent with a Class II deletion.^{56,57}

An atypical case involving the PWS/AS region was also identified (Case 3595). The proband presented with developmental delay, hypotonia and other features suggestive of AS. However, previous testing for AS by methylation analysis and sequencing of *UBE3A* were normal. By aCGH, an intragenic ~ 45 kb deletion of the *UBE3A* gene was identified (Fig. 2, C), which was confirmed by quantitative polymerase chain reaction in another clinical laboratory (University of Chicago Genetics Laboratory, data not shown). The deletion was subsequently shown to be inherited from the phenotypically normal mother by array analysis (data not shown). Given the mother's normal phenotype, her deletion most likely arose on a paternally inherited chromosome 15. Because of the small size of this deletion, it would likely be missed by most BAC arrays.

Examples of two additional imbalances detected by the targeted coverage of the array are shown in Figure 2. Case 3745 was referred due to developmental delay and failure to thrive and previous sequencing of the *UBE3A* gene was normal. Array analysis revealed a gain of the short arm of chromosome 17 involving band p11.2 (Fig. 2, D). The duplicated region is ~ 2.1 Mb in size, including ~ 25 known genes, and is within the critical region for the Duplication 17p11.2 syndrome⁵⁸ (PTLS [MIM 610883]). The duplication in this individual is smaller than the common 3.7 Mb duplication, which is the reciprocal duplication of the Smith-Magenis syndrome (SMS [MIM 182290]) deletion, but it includes the critical region.

Another example of a duplication detected by the targeted coverage is Case 2911, a female infant referred due to a congenital craniofacial anomaly. Microarray analysis identified a 400 kb duplication on the long arm of the X chromosome at band q28 (Fig. 2, E). This genomic region contains ~ 12 known genes, one of which is the *MECP2* gene [MIM 300005]. Duplications of *MECP2* in men cause infantile hypotonia, recurrent respiratory infections and severe MR.⁵⁹ In women, duplications of *MECP2* are usually associated with a normal phenotype. However, a similar phenotype as that observed in men can be observed if the X-inactivation status of the two X chromosomes is skewed. Because this female is only an infant, she is too young to assess the relationship between this duplication

and her developmental status and the finding will be considered in the context of her clinical presentation during her development. However, this duplication does affect a clinically significant region and this female will be at risk for having affected male offspring.

Clinically significant imbalances detected by genome-wide coverage

Ten clinically significant abnormalities, including eight deletions and two duplications, were identified by the genome-wide backbone coverage of the array (Table 2). These cases constitute 30.3% of all abnormalities (10 of 33) and 4.7% of the entire patient cohort (10 of 211). All abnormal results were confirmed with custom-designed homebrew FISH probes using clones from the 32K BAC set. The size of the deletions ranged from 0.6 to 6.8 Mb, and the duplication sizes were 0.7 and 8.1 Mb. Examples of these cases are shown in Figure 3.

The largest interstitial deletion detected using the oligonucleotide array involved the long arm of chromosome 14 at bands q23.1 to q23.3 (Case 3882, Fig. 3, A). The patient was referred due to developmental delay, dysmorphic features, cleft lip and palate, and spherocytosis. Previous G-banding analysis was normal. The deleted region identified by array analysis was ~6.8 Mb in size and included approximately 33 known genes. One of the genes in this region, *SPTB* (Spectrin beta chain, erythrocyte), causes autosomal dominant Spherocytosis Type I (SPH1 [MIM 182870]), consistent with the spherocytosis observed in the patient. On retrospective review of the G-banded chromosomes and with knowledge of the specific imbalance, no obvious abnormality of 14q was appreciated, demonstrating that even pure, relatively large deletions of genomic material can be difficult to identify by routine chromosome analysis.

Three additional interstitial deletions of ~5 Mb in size were detected. In Case 3841, microarray analysis demonstrated a deletion of the long arm of chromosome 2 involving band q31.1 (data not shown). The size of the deletion was determined to be ~4.8 Mb and included approximately 28 known genes. In Case 3484, results revealed an interstitial deletion of the long arm of chromosome 12 involving bands q21.2–q21.31 (Fig. 3, B). The deletion encompassed ~4.7 Mb and included ~12 known genes. This deletion was shown to be maternally inherited, but the mother was not phenotypically normal and shared similar abnormal clinical features with the proband. In Case 3715, aCGH results showed a deletion of the long arm of chromosome 15 involving bands q21.3 to q22.2. The deleted region was ~4.5 Mb in size (Fig. 3, C) and included approximately 30 known genes.

Of the clinically significant abnormalities, four cases were identified that were smaller than 5 Mb in size. Case 2818 was referred for developmental delay, failure to thrive, dysmorphic features, and a polycystic kidney. Array analysis revealed an interstitial loss of ~2.7 Mb on the short arm of chromosome 2 involving bands p15–p16.1 (Fig. 3, D), which was determined to be an interstitial deletion by confirmatory FISH analysis. The deleted region contains approximately nine known genes.

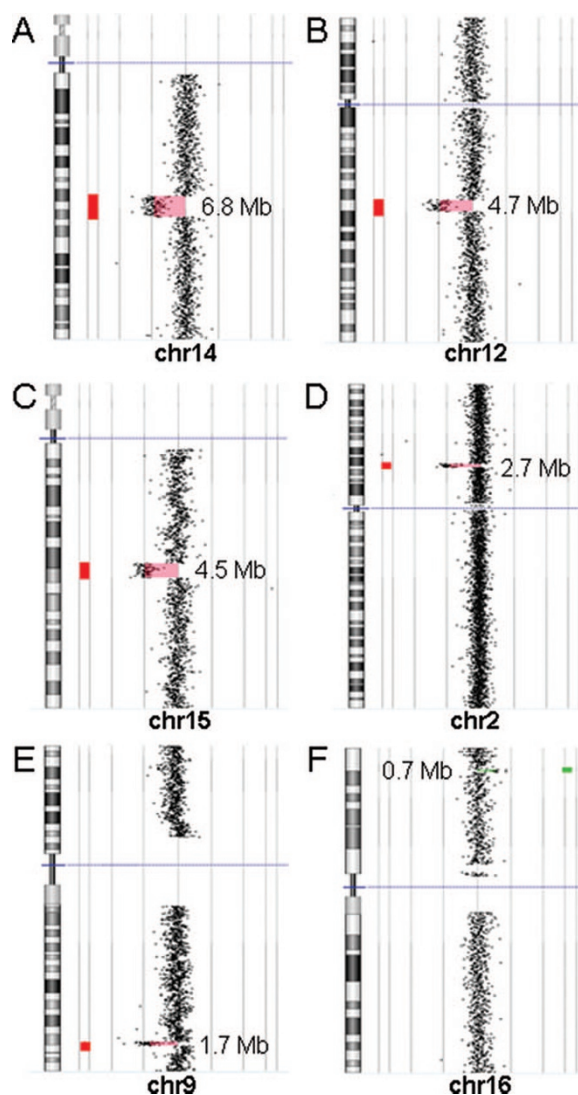


Fig. 3. Representative examples of interstitial imbalances detected by the genome-wide coverage of the custom microarray. A, Case 3882 with a 6.8 Mb deletion on the long arm of chromosome 14; B, Case 3484 with a 4.7 Mb deletion on the long arm of chromosome 12; C, Case 3715 with a 4.5 Mb deletion of the long arm of chromosome 15; D, Case 2818 with a 2.7 Mb deletion of the short arm of chromosome 2; E, Case 4153 with a 1.7 Mb deletion of the long arm of chromosome 9; F, Case 3596 with a 700 kb duplication involving the *A2BP1* gene on the short arm of chromosome 16.

A literature review revealed a previous report of two individuals with larger, overlapping deletions involving 2p15–16.1 which exhibited many of the same features as that observed in our patient, including developmental delay, similar dysmorphic features (as compared with the published pictures of the patients) and a multicystic kidney.⁶⁰

Case 4154 was referred for developmental delay and expressive language disorder. Microarray results revealed a deletion of the long arm of chromosome 15 involving bands q13.2–q13.3 (data not shown). The deleted region is ~2.2 Mb and includes ~6 known genes. Another small deletion was identified in a patient (Case 4153) who had bilateral strabismus and club feet at birth and infantile spasms which dissipated around age 5. She is now 23 years old and nonverbal with mild dys-

morphic features, seizures, aggressive behavior, and autistic disorder. Microarray analysis showed a deletion of the long arm of chromosome 9 involving bands q33.3–q34.11 (Fig. 3, E). The deleted region is ~1.7 Mb and includes approximately 10 known genes.

The smallest deletion detected by the genome-wide backbone coverage was identified in Case 3963. The individual was referred for MR, obesity, and behavior problems. The array results identified a deletion of the short arm of chromosome 16 involving band p11.2 (data not shown). The deleted region is ~598 kb, includes approximately 24 known genes. This genomic imbalance overlaps a previously reported microdeletion that was initially observed in a single family.⁶¹ Features reported in this family include aortic valve abnormalities, seizures and mild MR. The only known overlapping feature present in our patient is MR; she has no overt clinical evidence of a cardiac abnormality or seizure disorder, however an electroencephalogram and echocardiogram are currently pending. More recently, two reports have demonstrated that this microdeletion is a potentially frequent imbalance associated with a high risk of autism.^{62,63}

Two interstitial duplications were also detected by the genome-wide coverage of the array. In Case 2026, a gain involving the long arm of chromosome 2 was detected (data not shown). The size of the gain of 2q was determined to be ~8.1 Mb in size and included approximately 29 known genes. Subsequent FISH analysis confirmed the gain as an interstitial duplication. Previous G-banding analysis, performed in another laboratory, was reported as normal at the 750 band level and even in a retrospective analysis of G-banded chromosomes, the duplication was not easily appreciated.

Array analysis also revealed a small interstitial gain of the short arm of chromosome 16 involving band p13.3 in Case 3596 (Fig. 3, F). The gain of 16p was estimated to be ~0.7 Mb in size and included part of the *A2BP1/FOX1* gene [MIM 605104]. FISH analysis confirmed the gain as an interstitial duplication. Disruptions of the *A2BP1/FOX1* gene have previously been reported. A patient with autism, epilepsy, and global developmental delay was shown to have a partial deletion of this gene.⁶⁴ In addition, two unrelated patients with MR and epilepsy had translocation BPs that disrupted the *A2BP1/FOX1* gene.⁶⁵ The referring diagnosis for Case 3596 included microcephaly and features of Dandy Walker syndrome (DWS [MIM 220200]). Because the duplication present in this individual also involves the *A2BP1/FOX1* gene, it was predicted to be related to the abnormal clinical phenotype in this individual.

Common CNVs, benign familial variants and variants of uncertain clinical significance

Using a detection threshold of 500 kb, we identified 13 known CNVs that were >500 kb in size (Table 3). All of these are confirmed, common CNVs in the Database of Genomic Variants²⁷ that were present in at least 1% of the populations reported. As expected from our knowledge of the complex genomic architecture of the pericentromeric and subtelomeric

regions, ~46% of these large CNVs were located in these regions, many of which contained extensive segmental duplications.⁵⁰ Five CNVs were located in the pericentromeric regions, including 1q21.1, 2p11.2, 7q11.1, 10q11.22, and 15q11.2. One CNV was located in the 14q telomere region involving the *IGH* gene family. The remaining seven CNVs were located at interstitial regions (1p36.21, 4q13.2, 5q13.2, 8p23.1, 17q21.31, 19q13.2, and Xp22.31) which also were mainly comprised of segmental duplications representing high levels of shared sequence homologies with other regions of the genome. Of the 211 prospective cases, these 13 CNVs were detected in 42 cases (19.9%). A single CNV >500 kb in size was identified in 39 patients, two CNVs were detected in only two patients, and three CNVs were detected in only one patient.

Four (1.9%) cases were identified with previously published imbalances whose clinical significance remains uncertain (Table 3). Two of these cases involved the same genomic region, 16p13.11; a deletion (Case 3402) and a duplication (Case 3778) of this region were detected. At this time, the clinical significance of these findings is uncertain. Parental samples have been analyzed for the 16p13.11 deletion case and demonstrated that the deletion was paternally inherited. Imbalances of this region have been the topic of a recent publication that provides evidence for recurrent deletions and duplications of 16p13.11.⁶⁶ Deletions of this region were observed in individuals with MR, whereas duplications were observed in individuals with autism. However, it is important to note that both deletions and duplications of this region were also seen in some normal individuals in these families, but not in 600 normal control individuals examined.

The third case with an imbalance of uncertain clinical significance involves a duplication of the 1q21.1 region (3486). Recurrent deletions and duplications of this region have also been previously reported.^{23,32,67,68} Deletions of this region were observed in individuals with congenital cardiac defects, whereas duplications were observed in individuals with idiopathic MR or autism spectrum disorders. The deletions and duplications were observed in some normal parents, suggesting incomplete penetrance; however some carrier parents were found to have some subtle phenotypic abnormalities. Therefore, further studies of this region in normal control populations are needed to determine whether these imbalances are related to any specific phenotype or are coincidental findings.

The fourth imbalance, which involves 16p11.2, contained 24 genes, but was a duplication, rather than a deletion. Duplications of this region have recently been reported, however at this time, their clinical significance is not clear.⁶³

Only five cases (2.4%), not represented in the CNV database, required the analysis of parental samples before a clinical interpretation was possible. All five cases were shown to be inherited from a normal parent, and were therefore interpreted as most likely representing benign variants (Table 3). Four of these imbalances were duplications, whereas one was a deletion. Two duplications were located in targeted telomeric regions and were larger than our 50 kb threshold for these regions (120 and 330 kb). The other three imbalances were

Table 3
Common CNVs, benign familial variants and variants of uncertain clinical significance detected using the custom array

Chr location (case number)	Genomic coordinates ^a (NCBI build 35)	# Losses	# Gains	Total cases
Common CNVs^b				
Pericentromeric				
1q21.1	145,647,143 bp → 146,581,928 bp	2	0	2
2p11.2	87,236,959 bp → 87,844,680 bp	0	2	2
7q11.21	61,275,990 bp → 62,135,080 bp	1	1	2
10q11.22	46,404,919 bp → 47,125,152 bp	0	4	4
15q11.2	20,316,992 bp → 20,768,955 bp	1	1	2
Subtelomeric				
14q32.33	105,143,707 bp → 106,334,523 bp	2	1	3
Interstitial				
1p36.21	12,773,677 bp → 13,619,617 bp	0	1	1
4q13.2	69,231,557 bp → 69,789,502 bp	0	1	1
5q13.2	68,827,028 bp → 70,807,480 bp	5	1	6
8p23.1	6,793,320 bp → 8,391,131 bp	12	7	19
17q21.31	41,544,224 bp → 42,175,497 bp	1	1	2
19q13.2q13.31	47,934,635 bp → 48,433,150 bp	1	0	1
Xp22.31	6,348,457 bp → 7,907,247 bp	0	1	1
Benign familial variants				
1q41 (3099)	(214,883,459 bp → 215,523,361 bp)mat	1	0	1
5p15 (3591)	(9,623,606 bp → 10,322,569 bp)pat	0	1	1
11q22.1 (3590)	(99,418,604 bp → 100,132,465 bp)pat	0	1	1
11q25 (3923)	(133,886,842 bp → 134,216,882 bp)pat	0	1	1
20qter (2490)	(62,259,326 bp → 62,379,118 bp)mat	0	1	1
Uncertain clinical significance				
1q21.1 (3486)	(143,288,587 bp → 145,325,112 bp)	0	1	1
16p11.2 (3046)	(29,500,284 bp → 30,098,069 bp)mat	0	1	1
16p13.11 (3402)	(14,817,706 bp → 16,432,849 bp)pat	1	0	1
16p13.11 (3778)	(14,876,356 bp → 16,102,079 bp)	0	1	1

^aCoordinates indicate the largest region of the imbalance; smaller imbalances may exist in this copy number variant region.

^bCNVs were reported previously in the database of genomic variants.

detected by genome-wide coverage and were slightly larger than our 500 kb threshold (640–713 kb) and contain only a small number of genes (<3 genes per case).

DISCUSSION

In a series of 211 patients with unexplained MR/developmental delay, many of whom had previous cytogenetics studies showing a normal G-banded karyotype, clinically significant genomic imbalances were identified in 15.6% of cases. This frequency is consistent with other reports from genome-wide BAC-based aCGH^{23,69–71} and is higher than the typical yield from targeted arrays.^{17,18,72} Our results provide evidence that oligonucleotide-based aCGH is a reliable technology platform

in a clinical diagnostic setting for the accurate detection of single-copy gains and losses of genomic material, as previously suggested by several studies.^{39,40,73} Although intensity ratios of individual oligonucleotide probes are not a reliable indicator of copy number, identification of a series of contiguous probes (4–5) showing significant gain or loss has a very high predictive value. From our previous experience with BAC arrays,^{74,75} and current published literature, it seems that aCGH using long oligonucleotides (~60–85mers) has performance characteristics comparable with BAC arrays. However, although tiling-path BAC arrays have a theoretical resolution of ~50 kb,²⁴ these arrays with genome-wide coverage (>30,000 clones) are not trivial to produce. In contrast, genome-wide oligonucleotide arrays with hundreds of thousands of probes

can be efficiently produced and can easily reach a resolution of 50 kb or less.

In this study, all aCGH results were confirmed by FISH analysis using commercially available or “homebrew” probes from the 32K tiling-path BAC set. In our experience, and as suggested by published guidelines, FISH analysis is the preferred method for aCGH confirmation studies.^{3,55} In addition to validating the array result, FISH has the added advantage of demonstrating the mechanism for the imbalance. For example, a gain of genomic material detected by aCGH could be the result of a duplication or an unbalanced structural chromosome rearrangement, which would only be observed by analysis of metaphase chromosomes. FISH analysis is also the most cost effective and comprehensive method for parental follow-up studies, where other quantitative molecular methods would give a normal result when a balanced rearrangement (e.g., reciprocal or insertional translocation) is actually present in one of the parents. Identification of such balanced carriers has obvious implications for recurrence risk estimates and genetic counseling.

Accurate detection of targeted, clinically significant loci by oligonucleotide aCGH

For the targeted coverage of clinically relevant loci, an abnormality rate of 10.9% was observed (Table 1), consistent with other clinical targeted array formats primarily based on BAC clones.^{17,18} The most frequent of these abnormalities were telomere imbalances, which accounted for ~42% of the targeted abnormalities and 4.7% of the total imbalances found in this study. This data provides additional support for the prediction that telomere imbalances are overrepresented compared with other genomic regions.^{6,9,10} A recent study by our group examining the sizes and BPs of telomere imbalances showed that this overrepresentation is most likely due to random chromosome breakage and not to any specific molecular mechanism predisposing telomeres to more frequent breakage or rearrangements.⁴²

The coverage on the array at each telomere is equivalent to telomere FISH, but the added genome-wide resolution allows for simultaneous sizing of each imbalance in one assay, unlike telomere FISH studies or targeted arrays with limited telomeric coverage. It is important to note that the current genome-wide commercial arrays may not provide such equivalent coverage at all telomeres. An off-the-shelf commercial array, that is not specifically designed to cover clinically relevant regions, can miss telomere deletions or duplications that would have been detected by telomere FISH or targeted telomere coverage on custom designed arrays.⁴² Thus, it is important to carefully examine the content of any array to be used for telomere analysis.

Imbalances in the common microdeletion/microduplication regions accounted for ~6.6% of abnormalities. In the majority of cases, the BPs and extent of deletion/duplication were consistent with the common BPs mediated by nonallelic homologous recombination due to flanking segmental duplications with very high sequence homology.⁷⁶ An interesting exception to this was a 45 kb intragenic deletion of the *UBE3A*

gene in a patient clinically suspected to have AS. Previous methylation analysis and *UBE3A* sequencing studies were normal. Therefore, only targeted deletion analysis of this gene by oligonucleotide array provided a diagnosis. In this case, the deletion is smaller than an individual BAC clone and therefore below the level of resolution for most BAC arrays. Most current testing practices for AS do not include deletion analysis of *UBE3A*. Therefore, as demonstrated by this case, a category of patients with this syndrome is not being routinely tested for and gene deletion analysis should be added to diagnostic testing.

Two other microdeletion/duplication cases also exhibited atypical BPs. Case 3919 had a ~902 kb deletion involving the 22q11.2 deletion syndrome region which only included the most distal part of the typically deleted 3 Mb region. This region does not contain the smaller 1.5 Mb deleted region that has also been reported in individuals with 22q11.2 deletion syndrome. A recent article has reported a similar, but smaller, deletion in a female proband and her mother; the proband had obesity, hyperphasia and aggressive behavior and her mother had a major depressive disorder.⁷⁷ Our patient had a history of aortic aneurysm, a feature not reported in the other individuals with a similar, but smaller, deletion. Another atypical imbalance involved a 2.1 Mb duplication of 17p11.2, a region that causes Duplication 17p11.2 syndrome (Case 3745). The duplication in this patient is smaller than the common 3.7 Mb duplication, but includes the critical duplication region. Clinical findings associated with this syndrome include infantile hypotonia, failure to thrive, MR, autistic features, sleep apnea, and structural cardiovascular anomalies.⁵⁸ Our patient demonstrated several features noted in patients with the common deletion including developmental delay, failure to thrive, a structural cardiac anomaly, and severe speech delay. These three cases illustrate the clinical utility of high-density coverage within and surrounding targeted clinical regions on an array.

Increased yield of clinically significant imbalances provided by genome-wide coverage

In our series, the genome-wide backbone coverage identified an additional 10 cases (4.7%) of clinically significant abnormalities that would not be detected by current targeted array designs. The average size of the imbalances detected was ~3.7 Mb and contained an average of ~17 known genes (a range of 1–30) (Table 2). Most of these imbalances would be detected by current 1 Mb BAC arrays, which have been used in a number of clinical laboratories (usually following a normal targeted BAC array result). For the smaller imbalances, however, only a single clone may be involved and the exact sizing of the imbalance would not be achieved. Therefore, it may be more efficient and cost-effective to have both targeted and genome-wide coverage in a single assay.

Five of the 10 cases had imbalances ≥ 4.5 Mb, which would ordinarily be considered within or at the threshold of G-banded karyotype resolution. These included four deletions and one duplication. The clinical utility of accurately delineat-

ing the size of an imbalance and its gene content was immediately obvious in one case (Case 3882) in which one of the patient's clinical findings was spherocytosis. A de novo 14q23.1–q23.3 deletion of 6.8 Mb was identified by array analysis, which included the *SPTB* gene (spectrin beta chain, erythrocyte), mutations of which cause autosomal dominant Spherocytosis Type I. This case illustrates the power of genome-wide coverage in identifying deletions encompassing dominant Mendelian disease genes without the necessity to specifically “target” every dominant disease gene in the genome.

Three cases had somewhat smaller deletions, in the range of common microdeletion/duplication syndrome (between 1.5 and 3 Mb in size). The largest of these, a 2.7 Mb deletion of 2p (Case 2818), overlapped the deletions of two cases previously reported in the literature.⁶⁰ Because the phenotype of our patient was strikingly similar to that of the two published cases, the deletion was interpreted as pathogenic. The other two deletions in this size range also involved multiple genes: Case 4153 was a 1.7 Mb deletion of 9q that involved ~10 known genes and Case 4154 was a deletion of 2.2 Mb of 15q that involved ~6 known genes.

Only two cases identified by the genome-wide coverage were <1 Mb in size but considered to be clinically significant; a duplication of 16p13.3 (Case 3596) and a deletion of 16p11.2 (Case 3963). The duplication of 16p13.3 was ~700 kb in size and involved the *A2BP1/FOX1* gene, which has previously been implicated in MR, autism, and seizures.^{64,65} Because this patient was adopted, only one biological parent was available for analysis and did not show the same duplication.

The deletion of 16p11.2 was only ~600 kb in size, but included approximately 24 genes and is the same deletion recently reported in two studies of individuals with autism^{62,63} and overlaps part of the deleted 16p11.2 region reported by Ballif et al.⁶¹ Interestingly, this genomic region is flanked by segmental duplications with shared homology and this deletion seems to be a recurrent microdeletion associated with an increased risk for autism.

Identification and interpretation of CNVs and benign familial variants: a favorable cost-benefit ratio

Over the past several years, multiple studies have revealed that CNVs are very common in the human genome.^{27–34} Because of this normal variation, some authors have argued that genome-wide aCGH analysis is premature in a clinical setting because of our current lack of knowledge regarding the clinical significance of CNVs and voiced a concern that the majority of patients will require parental follow-up.^{43,44} However, the initial lack of knowledge regarding CNVs is rapidly being filled because of international efforts to produce data on normal individuals and populations in public databases such as the Database of Genomic Variants²⁷ and DECIPHER (<http://decipher.sanger.ac.uk/>).

Although the high frequency of submicroscopic copy number variation in the human genome is a recent and surprising discovery, the presence of variation in the human karyotype at a microscopically visible level has been known for a long time. Although most of this variation involves highly repeated DNA

sequences in the heterochromatic regions of the genome, there is an increasing awareness of microscopically visible variations in euchromatic regions.^{78,79} A small number of recurrent, or common, euchromatic variants have been reported, including those involving 8p23.1, 9p12, 9q12, 15q11.2, and 16p11.2.⁷⁸ There are also a growing number of private variants, involving single families, which have been reported in the Chromosome Anomaly Collection Database (<http://www.ngri.org.uk/Wessex/collection/index.htm>). Most of these euchromatic variants have not been characterized in detail at a molecular level, but suggest that there are large genomic regions that are relative gene deserts and/or do not contain genes which are sensitive to single copy dosage changes.

Although our size threshold of 500 kb greatly reduced the detection of CNVs, we still identified a number of CNVs in our patient population (Table 3). These included 13 common, previously reported CNVs, for which we did not perform any parental follow-up studies. A significant percentage of these common CNVs are in the pericentromeric and subtelomeric regions, which are known to be enriched for segmental duplications and show high rates of copy number variation. In the remaining euchromatic portion of the genome, we observed seven other common CNVs. The highest number of common CNVs identified in a single patient was two, with a range of 0–2 in our patient cohort.

To date, we have performed follow-up parental studies on 2.4% of our patients to assess the clinical significance. Our experience with this patient cohort supports Veltman and de Vries^{25,26} viewpoint that parental follow-up of ~3% is worth the diagnostic benefit of identifying an additional 5% of causative abnormalities.

These common CNVs and euchromatic variants could be removed from diagnostic arrays to make clinical interpretation easier. However, some of these regions are quite large and by removing them, aberrations could be missed that involve these and surrounding regions. Furthermore, hemizygous deletions or duplications of these regions may be benign, but homozygous deletions or duplications may not be normal (e.g., homozygous deletions of the juvenile nephronophthisis [MIM 256100] involving the *NPHP1* gene [MIM 607100] at 2q13).⁸⁰

aCGH as a molecular karyotype versus a Mendelian disease diagnostic chip

For the evaluation of a child with unexplained developmental delay/MR, with or without dysmorphic features or additional birth defects, the G-banded karyotype has been the primary diagnostic tool for over 30 years. For detection of deletions and duplications, aCGH clearly offers a more powerful and sensitive technology platform than G-banding. However, the ideal design of such arrays has not yet been determined. One trend, based on a targeted approach, has been to combine cytogenetic targets (telomeres, centromeres, microdeletion/duplication syndromes) with Mendelian disease gene targets. The latter are largely autosomal dominant diseases known to be due to haploinsufficiency (loss-of-function

mutations), with a variable percentage of causative mutations resulting from deletions.

However, as discussed previously, such a targeted approach leaves many large chromosomal gaps and cannot be considered a molecular karyotype. To address this issue, we chose to design a genome-wide array that would increase the resolution of G-banding to the equivalent of a 6000 band karyotype. For targeted, clinically relevant regions, including telomeres, centromeres, common microdeletion/microduplication syndromes and selected Mendelian disorders, the coverage is even more dense allowing for imbalances of <50 kb to be detected. This combined approach leverages the advantages of both targeted and genome-wide arrays in a single assay or diagnostic test.

The development of a custom array ensures that all clinically relevant genomic regions have adequate coverage for diagnostic testing. Many of the currently available commercial arrays have avoided the complex genomic regions near centromeres and telomeres that are clearly important for clinical diagnostics. Furthermore, coverage of the PAR1 and PAR2, which are important in identifying imbalances of the sex chromosomes, are not included on most commercial arrays.

With the rapidly progressing technology of oligonucleotide array printing, the potential exists to increase the number of probes to millions in the next several years. Ongoing discussions will need to evaluate whether the evolving strategy should be to continue to add targeted coverage for every known autosomal dominant loss of function gene or to use a genome-wide high-resolution cytogenetic approach. Gene specific coverage will need to be carefully assessed because high density coverage for each gene is needed to accurately test for intragenic deletions or duplications; this level of detail may be better covered in gene specific applications rather than genome-wide analyses.

Can aCGH replace the G-banded karyotype as a primary diagnostic tool?

To date, aCGH has been recommended as an adjunctive test to the G-banded karyotype until sufficient data are obtained to evaluate its potential as a primary test for children with unexplained MR/developmental delay. Significant data are accumulating that targeted arrays have an abnormal yield of 5–10% after a normal G-banded karyotype. Genome-wide arrays identify an additional ~5% abnormalities not detected by targeted arrays or G-banding. If genome-wide arrays are performed before G-banding, they would detect all imbalances detected by G-banding plus an additional 10–15% not currently detected. Furthermore, a recent study demonstrated that the use of aCGH actually costs less per diagnosis than karyotype analysis because earlier diagnoses save the costs of additional diagnostic tests.⁸¹

The benefit of using genome-wide aCGH as a primary diagnostic tool versus the concern over the detection of CNVs is also becoming clear. Data on CNVs in normal populations is rapidly accumulating, mainly from research studies, and is being deposited in an international database.²⁷ It would be similarly valuable to collect clinical aCGH data in a shared data-

base to continue to increase our understanding of the regions of the genome that cause significant developmental effects when copy number changes are present. The use of a common array format by a consortium of laboratories would improve this data collection, because as previously pointed out, the use of data from a variety of methodologies and arrays to populate CNV databases makes it difficult to accurately interpret results and apply in a clinical setting.⁷² This effort would allow the development of a Human Gene Dosage Map for constitutional copy number changes which would not only aid in diagnostic testing, but would also enhance our understanding of the functional significance of genes and other sequences in the human genome.

ACKNOWLEDGMENTS

The work was funded in part by NIH Grant RO1 MH074090 (to D.H.L. and C.L.M.).

The authors thank Elijah Wallace, MS for expert technical assistance, David Chrimes, PhD (BlueGnome Ltd.) for assistance with the Blue Fuse software, Lars Feuk, PhD (The Hospital for Sick Children, Toronto) for assistance with the CNV analysis and Condie Carmack, PhD and Nick Sampas, PhD (Agilent Technologies) for assistance with the array design. We also thank our laboratory genetic counselors, Vanessa Rangel Miller, MS and Nikki Justice, MS for assistance with the array cases, and the clinicians, genetic counselors and families involved in these studies.

References

- Shaffer LG, Tommerup N, editors. ISCN 2005. An international system for human cytogenetics nomenclature. Basel: Karger, 2005.
- Shevell M, Ashwal S, Donley D, Flint J, et al. Practice parameter: evaluation of the child with global developmental delay: report of the Quality Standards Subcommittee of the American Academy of Neurology and The Practice Committee of the Child Neurology Society. *Neurology* 2003;60:367–380.
- Shaffer LG. American College of Medical Genetics guideline on the cytogenetic evaluation of the individual with developmental delay or mental retardation. *Genet Med* 2005;7:650–654.
- Yunis JJ, Sawyer JR, Ball DW. The characterization of high-resolution G-banded chromosomes of man. *Chromosoma* 1978;67:293–307.
- Yunis JJ. Mid-prophase human chromosomes. The attainment of 2000 bands. *Hum Genet* 1981;56:293–298.
- Ledbetter SA, Kuwano A, Dobyns WB, Ledbetter DH. Microdeletions of chromosome 17p13 as a cause of isolated lissencephaly. *Am J Hum Genet* 1992;50:182–189.
- Kuwano A, Mutirangura A, Dittrich B, Buiting K, et al. Molecular dissection of the Prader-Willi/Angelman syndrome region (15q11–13) by YAC cloning and FISH analysis. *Hum Mol Genet* 1992;1:417–425.
- Desmaze C, Scambler P, Prieur M, Halford S, et al. Routine diagnosis of DiGeorge syndrome by fluorescent in situ hybridization. *Hum Genet* 1993;90:663–665.
- Ning Y, Rosenberg M, Biesecker LG, Ledbetter DH. Isolation of the human chromosome 22q telomere and its application to detection of cryptic chromosomal abnormalities. *Hum Genet* 1996;97:765–769.
- Knight SJ, Lese CM, Precht KS, Kuc J, et al. An optimized set of human telomere clones for studying telomere integrity and architecture. *Am J Hum Genet* 2000;67:320–332.
- Biesecker LG. The end of the beginning of chromosome ends. *Am J Med Genet* 2002;107:263–266.
- Ravnan JB, Tepperberg JH, Papenhausen P, Lamb AN, 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.
- Kallioniemi A, Kallioniemi OP, Sudar D, Rutovitz D, et al. Comparative genomic hybridization for molecular cytogenetic analysis of solid tumors. *Science* 1992;258:818–821.

14. Pinkel D, Seagraves R, Sudar D, Clark S, et al. High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays. *Nat Genet* 1998;20:207–211.
15. Pinkel D, Albertson DG. Array comparative genomic hybridization and its applications in cancer. *Nat Genet* 2005;37:S11–S17.
16. Vissers LE, de Vries BB, Osoegawa K, Janssen IM, et al. Array-based comparative genomic hybridization for the genomewide detection of submicroscopic chromosomal abnormalities. *Am J Hum Genet* 2003;73:1261–1270.
17. Shaffer LG, Kashork CD, Saleki R, Rorem E, et al. Targeted genomic microarray analysis for identification of chromosome abnormalities in 1500 consecutive clinical cases. *J Pediatr* 2006;149:98–102.
18. Lu X, Shaw CA, Patel A, Li J, et al. Clinical implementation of chromosomal microarray analysis: summary of 2513 postnatal cases. *PLoS ONE* 2007;2:e327.
19. Snijders AM, Nowak N, Seagraves R, Blackwood S, et al. Assembly of microarrays for genome-wide measurement of DNA copy number. *Nat Genet* 2001;29:263–264.
20. Fiegler H, Carr P, Douglas EJ, Burford DC, et al. DNA microarrays for comparative genomic hybridization based on DOP-PCR amplification of BAC and PAC clones. *Genes Chromosomes Cancer* 2003;36:361–374.
21. Ishkanian AS, Malloff CA, Watson SK, DeLeeuw RJ, et al. A tiling resolution DNA microarray with complete coverage of the human genome. *Nat Genet* 2004;36:299–303.
22. Krzywinski M, Bosdet I, Smalilou D, Chiu R, et al. A set of BAC clones spanning the human genome. *Nucleic Acids Res* 2004;32:3651–3660.
23. de Vries BB, Pfundt R, Leisink M, Koolen DA, et al. Diagnostic genome profiling in mental retardation. *Am J Hum Genet* 2005;77:606–616.
24. Coe BP, Ylstra B, Carvalho B, Meijer GA, et al. Resolving the resolution of array CGH. *Genomics* 2007;89:647–653.
25. Veltman JA, de Vries BB. Diagnostic genome profiling: unbiased whole genome or targeted analysis? *J Mol Diagn* 2006;8:534–537; discussion 537–539.
26. Veltman JA, de Vries BB. Whole-genome array comparative genome hybridization: the preferred diagnostic choice in postnatal clinical cytogenetics. *J Mol Diagn* 2007;9:277.
27. Iafrate AJ, Feuk L, Rivera MN, Listewnik ML, et al. Detection of large-scale variation in the human genome. *Nat Genet* 2004;36:949–951.
28. Sebat J, Lakshmi B, Troge J, Alexander J, et al. Large-scale copy number polymorphism in the human genome. *Science* 2004;305:525–528.
29. Sharp AJ, Locke DP, McGrath SD, Cheng Z, et al. Segmental duplications and copy-number variation in the human genome. *Am J Hum Genet* 2005;77:78–88.
30. Tuzun E, Sharp AJ, Bailey JA, Kaul R, et al. Fine-scale structural variation of the human genome. *Nat Genet* 2005;37:727–732.
31. Conrad DF, Andrews TD, Carter NP, Hurler ME, et al. A high-resolution survey of deletion polymorphism in the human genome. *Nat Genet* 2006;38:75–81.
32. Redon R, Ishikawa S, Fitch KR, Feuk L, et al. Global variation in copy number in the human genome. *Nature* 2006;444:444–454.
33. Wong KK, deLeeuw RJ, Dosanjh NS, Kimm LR, et al. A comprehensive analysis of common copy-number variations in the human genome. *Am J Hum Genet* 2007;80:91–104.
34. Simon-Sanchez J, Scholz S, Fung HC, Matarin M, et al. Genome-wide SNP assay reveals structural genomic variation, extended homozygosity and cell-line induced alterations in normal individuals. *Hum Mol Genet* 2007;16:1–14.
35. Ylstra B, van den Ijssel P, Carvalho B, Brakenhoff RH, et al. BAC to the future! or oligonucleotides: a perspective for micro array comparative genomic hybridization (array CGH). *Nucleic Acids Res* 2006;34:445–450.
36. Hehir-Kwa JY, Egmont-Petersen M, Janssen IM, Smeets D, et al. Genome-wide copy number profiling on high-density bacterial artificial chromosomes, single-nucleotide polymorphisms, and oligonucleotide microarrays: a platform comparison based on statistical power analysis. *DNA Res* 2007;14:1–11.
37. Barrett MT, Scheffer A, Ben-Dor A, Sampas N, et al. Comparative genomic hybridization using oligonucleotide microarrays and total genomic DNA. *Proc Natl Acad Sci U S A* 2004;101:17765–17770.
38. Selzer RR, Richmond TA, Pofahl NJ, Green RD, et al. Analysis of chromosome breakpoints in neuroblastoma at sub-kilobase resolution using fine-tiling oligonucleotide array CGH. *Genes Chromosomes Cancer* 2005;44:305–319.
39. Jacquemont ML, Sanlaville D, Redon R, Raoul O, et al. Array-based comparative genomic hybridisation identifies high frequency of cryptic chromosomal rearrangements in patients with syndromic autism spectrum disorders. *J Med Genet* 2006;43:843–849.
40. Aradhya S, Manning MA, Splendore A, Cherry AM. Whole-genome array-CGH identifies novel contiguous gene deletions and duplications associated with developmental delay, mental retardation, and dysmorphic features. *Am J Med Genet A* 2007;143:1431–1441.
41. Toruner GA, Streck DL, Schwab MN, Dermody JJ. An oligonucleotide based array-CGH system for detection of genome wide copy number changes including subtelomeric regions for genetic evaluation of mental retardation. *Am J Med Genet A* 2007;143:824–829.
42. Martin CL, Nawaz N, Baldwin EL, Wallace EJ, et al. The evolution of molecular ruler analysis for characterizing telomere imbalances: from FISH to array CGH. *Genet Med* 2007;9:566–573.
43. Bejjani BA, Shaffer LG. Application of array-based comparative genomic hybridization to clinical diagnostics. *J Mol Diagn* 2006;8:528–533.
44. Bejjani BA, Shaffer LG. Targeted array CGH. *J Mol Diagn* 2006;8:537–539.
45. Martin CL, Waggoner DJ, Wong A, Uhrig S, et al. “Molecular rulers” for calibrating phenotypic effects of telomere imbalance. *J Med Genet* 2002;39:734–740.
46. Daniels RJ, Peden JF, Lloyd C, Horsley SW, et al. Sequence, structure and pathology of the fully annotated terminal 2 Mb of the short arm of human chromosome 16. *Hum Mol Genet* 2001;10:339–352.
47. Kent WJ, Sugnet CW, Furey TS, Roskin KM, et al. The human genome browser at UCSC. *Genome Res* 2002;12:996–1006.
48. Karolchik D, Baertsch R, Diekhans M, Furey TS, et al. The UCSC genome browser database. *Nucleic Acids Res* 2003;31:51–54.
49. Baldwin EL, May LF, Justice AN, Martin CL, et al. Mechanisms and consequences of small supernumerary marker (ring) chromosomes: from Barbara McClintock to modern genetic counseling issues. *Am J Hum Genet* 2008;82:398–410.
50. Bailey JA, Yavor AM, Massa HF, Trask BJ, et al. Segmental duplications: organization and impact within the current human genome project assembly. *Genome Res* 2001;11:1005–1017.
51. Lander ES, Linton LM, Birren B, Nusbaum C, et al. Initial sequencing and analysis of the human genome. *Nature* 2001;409:860–921.
52. Furey TS, Haussler D. Integration of the cytogenetic map with the draft human genome sequence. *Hum Mol Genet* 2003;12:1037–1044.
53. Hsu F, Kent WJ, Clawson H, Kuhn RM, et al. The UCSC known genes. *Bioinformatics* 2006;22:1036–1046.
54. Chong SS, Pack SD, Roschke AV, Tanigami A, et al. A revision of the lissencephaly and Miller-Dieker syndrome critical regions in chromosome 17p13.3. *Hum Mol Genet* 1997;6:147–155.
55. Vermeesch JR, Fiegler H, de Leeuw N, Szuhai K, et al. Guidelines for molecular karyotyping in constitutional genetic diagnosis. *Eur J Hum Genet* 2007;15:1105–1114.
56. Christian SL, Robinson WP, Huang B, Mutirangura A, et al. Molecular characterization of two proximal deletion breakpoint regions in both Prader-Willi and Angelman syndrome patients. *Am J Hum Genet* 1995;57:40–48.
57. Fantes JA, Mewborn SK, Lese CM, Hedrick J, et al. Organisation of the pericentromeric region of chromosome 15: at least four partial gene copies are amplified in patients with a proximal duplication of 15q. *J Med Genet* 2002;39:170–177.
58. Potocki L, Bi W, Treadwell-Deering D, Carvalho CM, et al. Characterization of Potocki-Lupski syndrome (dup(17)(p11.2p11.2)) and delineation of a dosage-sensitive critical interval that can convey an autism phenotype. *Am J Hum Genet* 2007;80:633–649.
59. Villard L. MECP2 mutations in males. *J Med Genet* 2007;44:417–423.
60. Rajcan-Separovic E, Harvard C, Liu X, McGillivray B, et al. Clinical and molecular cytogenetic characterisation of a newly recognised microdeletion syndrome involving 2p15–16.1. *J Med Genet* 2007;44:269–276.
61. Ballif BC, Hornor SA, Jenkins E, Madan-Khetarpal S, et al. Discovery of a previously unrecognized microdeletion syndrome of 16p11.2-p12.2. *Nat Genet* 2007;39:1071–1073.
62. Kumar RA, KaraMohamed S, Sudi J, Conrad DF, et al. Recurrent 16p11.2 microdeletions in autism. *Hum Mol Genet* 2008;17:628–638.
63. Weiss LA, Shen Y, Korn JM, Arking DE, et al. Association between Microdeletion and Microduplication at 16p11.2 and Autism. *N Engl J Med* 2008;358:667–675.
64. Martin CL, Duvall JA, Ilkin Y, Simon JS, et al. Cytogenetic and molecular characterization of A2BP1/FOX1 as a candidate gene for autism. *Am J Med Genet B Neuropsychiatr Genet* 2007;144:869–876.
65. Bhalla K, Phillips HA, Crawford J, McKenzie OL, et al. The de novo chromosome 16 translocations of two patients with abnormal phenotypes (mental retardation and epilepsy) disrupt the A2BP1 gene. *J Hum Genet* 2004;49:308–311.
66. Ullmann R, Turner G, Kirchoff M, Chen W, et al. Array CGH identifies reciprocal 16p13.1 duplications and deletions that predispose to autism and/or mental retardation. *Hum Mutat* 2007;28:674–682.
67. Christiansen J, Dyck JD, Elyas BG, Lilley M, et al. Chromosome 1q21.1 contiguous gene deletion is associated with congenital heart disease. *Circ Res* 2004;94:1429–1435.
68. Sharp AJ, Hansen S, Selzer RR, Cheng Z, et al. Discovery of previously unidentified genomic disorders from the duplication architecture of the human genome. *Nat Genet* 2006;38:1038–1042.
69. Shaw-Smith C, Redon R, Rickman L, Rio M, et al. Microarray based comparative genomic hybridisation (array-CGH) detects submicroscopic chromosomal deletions and duplications in patients with learning disability/mental retardation and dysmorphic features. *J Med Genet* 2004;41:241–248.
70. Menten B, Maas N, Thienpont B, Buysse K, et al. Emerging patterns of cryptic

- chromosomal imbalance in patients with idiopathic mental retardation and multiple congenital anomalies: a new series of 140 patients and review of published reports. *J Med Genet* 2006;43:625–633.
71. Rosenberg C, Knijnenburg J, Bakker E, Vianna-Morgante AM, et al. Array-CGH detection of micro rearrangements in mentally retarded individuals: clinical significance of imbalances present both in affected children and normal parents. *J Med Genet* 2006;43:180–186.
 72. Shaffer LG, Bejjani BA, Torchia B, Kirkpatrick S, et al. The identification of microdeletion syndromes and other chromosome abnormalities: cytogenetic methods of the past, new technologies for the future. *Am J Med Genet C Semin Med Genet* 2007;145:335–345.
 73. Fan YS, Jayakar P, Zhu H, Barbouth D, et al. Detection of pathogenic gene copy number variations in patients with mental retardation by genomewide oligonucleotide array comparative genomic hybridization. *Hum Mutat* 2007;28:1124–1132.
 74. Schaeffer AJ, Chung J, Heretis K, Wong A, et al. Comparative genomic hybridization-array analysis enhances the detection of aneuploidies and submicroscopic imbalances in spontaneous miscarriages. *Am J Hum Genet* 2004;74:1168–1174.
 75. Wong A, Lese Martin C, Heretis K, Ruffalo T, et al. Detection and calibration of microdeletions and microduplications by array-based comparative genomic hybridization and its applicability to clinical genetic testing. *Genet Med* 2005;7:264–271.
 76. Shaw CJ, Lupski JR. Implications of human genome architecture for rearrangement-based disorders: the genomic basis of disease. *Hum Mol Genet* 2004;13 Spec No 1:R57–R64.
 77. D'Angelo CS, Jehee FS, Koiffmann CP. An inherited atypical 1 Mb 22q11.2 deletion within the DGS/VCFS 3 Mb region in a child with obesity and aggressive behavior. *Am J Med Genet A* 2007;143:1928–1932.
 78. Barber JC. Directly transmitted unbalanced chromosome abnormalities and eukaryotic variants. *J Med Genet* 2005;42:609–629.
 79. Daniel A, Darmanian A, Peters G, Goodwin L, et al. An innocuous duplication of 11.2 Mb at 13q21 is gene poor: sub-bands of gene paucity and pervasive CNV characterize the chromosome anomalies. *Am J Med Genet A* 2007;143:2452–2459.
 80. Saunier S, Calado J, Benessy F, Silbermann F, et al. Characterization of the NPHP1 locus: mutational mechanism involved in deletions in familial juvenile nephronophthisis. *Am J Hum Genet* 2000;66:778–789.
 81. Wordsworth S, Buchanan J, Regan R, Davison V, et al. Diagnosing idiopathic learning disability: a cost-effectiveness analysis of microarray technology in the National Health Service of the United Kingdom. *Genomic Med* 2007;1:35–45.