npg

ARTICLE

Custom oligonucleotide array-based CGH: a reliable diagnostic tool for detection of exonic copy-number changes in multiple targeted genes

Aurélie Vasson[1], Céline Leroux[1], Lucie Orhant[1], Mathieu Boimard[1], Aurélie Toussaint[1], Chrystel Leroy[1], Virginie Commere[1], Tiffany Ghiotti[1], Nathalie Deburgrave[1], Yoann Saillour[2], Isabelle Atlan[1], Corinne Fouveaut[1], Cherif Beldjord[1,2], Sophie Valleix[1,2], France Leturcq[1,2], Catherine Dodé[1,2], Thierry Bienvenu[1,2], Jamel Chelly[1,2] and Mireille Cossée*[,1,2]

The frequency of disease-related large rearrangements (referred to as copy-number mutations, CNMs) varies among genes, and search for these mutations has an important place in diagnostic strategies. In recent years, CGH method using custom-designed high-density oligonucleotide-based arrays allowed the development of a powerful tool for detection of alterations at the level of exons and made it possible to provide flexibility through the possibility of modeling chips. The aim of our study was to test custom-designed oligonucleotide CGH array in a diagnostic laboratory setting that analyses several genes involved in various genetic diseases, and to compare it with conventional strategies. To this end, we designed a 12-plex CGH array (135k; 135 000 probes/subarray) (Roche Nimblegen) with exonic and intronic oligonucleotide probes covering 26 genes routinely analyzed in the laboratory. We tested control samples with known CNMs and patients for whom genetic causes underlying their disorders were unknown. The contribution of this technique is undeniable. Indeed, it appeared reproducible, reliable and sensitive enough to detect heterozygous single-exon deletions or duplications, complex rearrangements and somatic mosaicism. In addition, it improves reliability of CNM detection and allows determination of boundaries precisely enough to direct targeted sequencing of breakpoints. All of these points, associated with the possibility of a simultaneous analysis of several genes and scalability 'homemade' make it a valuable tool as a new diagnostic approach of CNMs.
European Journal of Human Genetics (2013) 21, 977–987; doi:10.1038/ejhg.2012.279; published online 23 January 2013

Keywords: CGH array; oligonucleotide probes; copy-number mutations; CNMs; complex rearrangements

INTRODUCTION

Large intragenic deletions and duplications are frequent causes of genetic disorders. The term of copy-number mutations (CNMs) is used, from the standpoint pathological relevance, in distinction to the neutral term 'copy-number variation' CNV, in accordance with previous studies.[1] Thus, in the diagnostic context, search for CNMs represents a major purpose, particularly if they are frequent such as for Duchenne or Becker dystrophies (DMD/BMD) in which they represent 60–70% of mutations.[2,3] However, in diseases for which CNMs are rare, their search is important for diagnostic confirmation if screening for point mutation is negative or incompletely conclusive. For example, in cystic fibrosis (CF) CNMs represent 1.5% of known *CFTR* mutations and are searched for in patients heterozygous for point mutations in *CFTR*.[4]

Several scanning methods for CNMs, including Southern blotting, multiplex amplifiable probe hybridization (MAPH), quantitative multiplex PCR of short fluorescent fragments (QMPSF), multiplex ligation-dependent probe amplification (MLPA) are currently used[5–7] but they have several disadvantages. The preparation of high-quality Southern blot is technically demanding, is time consuming and its

sensitivity is generally low and limited to very large CNMs.[8] It has been shown that QMPSF approach used in several studies was inherently biased in favor of the detection of deletions over duplications, suggesting that a change of copy-number from two to one (in the case of heterozygous deletions) is more readily identifiable than change from two to three (in the case of heterozygous duplications).[4,9] MLPA approach seems more reliable as it is based on hybridization of probes to genomic regions of interest. However, this technique is available in commercial kit forms in which the number of probes is usually limited to one per exon. For diseases with rare CNMs, commercials kits such as MLPA are not available and scanning methods are represented by 'home made' techniques such as semi-quantitative-fluorescent-PCR (QF-PCR) and real-time quantitative-PCR (Q-PCR) that are time consuming for implementation. As PCR-based techniques and MLPA are restricted to a limited number of targeted sequences, they can fail to detect some rearrangements; conversely false-positive single-exons losses can result from single nucleotide polymorphisms (SNPs) affecting primer or probe sequences.[10] Finally, these scanning methods do not allow determination of CNM boundaries, and a limited number of

genes can be concomitantly analyzed. In general, each gene requires its specific kit or 'home-made' technique for CNM detection.

The recent emergence of array comparative genomic hybridization has revolutionized the ability to identify CNMs associated with various diseases. This approach was first used for detecting large CNMs at the scale of multiple contiguous genes in whole genome analysis.[11] In recent years, the development of oligonucleotide probes for hybridization on microarrays allowed to explore smaller CNMs at the scale of exons. Single-locus arrays were first validated to detect exonic and intronic CNMs within the *DMD* locus and the *CFTR* locus in patients suspected of having mutations in these genes.[1,8,10,12] Thereafter, several teams (including our) implemented multiple gene CGH arrays.[13–15] Custom oligonucleotide CGH array emerged then as a powerful tool for high-resolution detection of genomic CNMs, with the flexibility provided through customized array designs. The recent possibility of increasing the density of probes loaded on chips allowed the emergence of high-density (HD) chips arrays thereby increasing the number of genes tested.[16]

In our laboratory of Molecular Biology, Cochin Hospital (Paris), we are performing molecular diagnosis of several genetic diseases with various clinical aspects, mutational spectrum and modes of inheritance. Some of them are known to be caused, in different proportions, to CNMs (Table 1).[1–4,17–33] Conventional methods used to search for CNMs are mainly based on QF-PCR and MLPA. In order to replace the time-consuming current scanning methods by a unique technique for CNMs detection, we first developed a 72k four-plex array covering the 158 exons of eight disease-related genes: *DMD*, sarcoglycan genes and *CFTR*.[13] Exonic copy-number changes were identified with a high resolution, as abnormalities of about 1.5–2 kb could be detected, as well as a mosaic deletion.[13] We then wanted to implement this approach to the totality of genes tested in our laboratory that are known or suspected to be prone to CNMs (26 genes). We took advantage of the advent of HD chips to develop a custom-targeted 12-plex 135K oligonucleotide-based CGH array covering the 344 exons of these genes. We report implementation and validation of this array and its applications in a diagnostic setting to test different diseases with great clinical and genetic heterogeneity.

PATIENTS AND METHODS
Patients and control samples
The patients were referred to the Laboratory of Molecular Biology, Cochin Hospital, for molecular analysis of the genes of interest. Reference DNAs were obtained from control patients with a well characterized deletion of exons 7 and 8 of the *CFTR* gene.[13] Genomic DNAs were extracted from leukocytes using standard procedures (phenol extraction or Wizard Genomic DNA Isolation System, Promega, Madison, WI, USA).

CGH array
Microarray construction. We devised, with Roche Nimblegen support, 12-plex oligonucleotide-based CGH arrays to explore the whole 26 genes including promoters. The design was made by taking in consideration size and constitution of each gene (Table 2). Using the data from the human genome (http://genome.ucsc.edu/cgi-bin/hgGateway, NCBI36/hg18), 135 000 oligonucleotide probes covering all genomic regions plus 2 kb at each extremity of the 26 genes were designed for each subarray. On the basis of experimental results, different average tiling intervals (ie, spacing between 5′ ends of probes) were selected (Table 2). 'Backbone' probes covering chromosomes corresponding to the genes of interest were added at a lower density (spacing of 20–25 kb) in intergenic regions. All probes have similar characteristics: isothermal probes with melting temperature (Tm) of 76 °C and average probe length 60 bases. To avoid cross-hybridization, all probes were compared with the entire hg18 genome using Basic Local Alignment and Search Toll. Any probe that did not

map uniquely was removed except those in the pseudoautosomal regions on chromosome X and Y for which two locations were tolerated. Roche NimbleGen manufactured the array (www.nimblegen.com). Sequences of the 135 000 probes are available on request.

Fluorescent DNA labeling, microarray hybridization. DNA concentration and quality were evaluated by NanoDrop and agarose gel migration. The reference DNA used for each patient's DNA was extracted using the same technique. We did sex match between each sample and reference DNA. Each DNA sample (1 μg) was labeled using a NimbleGen Dual-Color DNA labeling Kit as described previously[13] according to the manufacturer's protocol (Roche NimbleGen). After denaturation, hybridization was carried out on a NimbleGen Hybridization System for 40 h at 42 °C. The array was then washed by using NimbleGen Wash System (Roche NimbleGen), dried by centrifugation and scanned at 2 μm resolution by using InnoScann900 scanner (Innopsys, Toulouse, France).

Data analysis. Fluorescence intensity raw data were obtained from scanned images of the oligonucleotide arrays by using NimbleScan 2.6 extraction software (Roche NimbleGen). For each spot on the array, $\log_2$ ratio of the Cy3-labeled test sample *versus* Cy-5 reference sample was calculated and visualization of the results was obtained using the Signal map software (Roche NimbleGen). Quality of the experiment was ascertained by the mad1.dr value (medium absolute deviation of the $\log_2$ ratio difference between consecutive probes) that provides a surrogate measure of experimental noise and should be <0.23.

DNA sequencing
CGH array results were used to target the genomic region to design primers for sequencing breakpoints. Primers were first selected within a 500-bp interval and further in a 1-kb interval if necessary. Oligonucleotide primer pairs were designed with the help of the Primer3Plus online tool (http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi). Sequences of primers are available on request. Bidirectional sequencing of the purified PCR products was performed on an Applied (3130XT) automated capillary sequencer (Applied Biosystems, Foster City, CA, USA) (protocols available on request).

RESULTS
Development of a custom 12-plex CGH array for CNM detection in 26 genes
In order to define the CGH array design that allows analysis of the greatest number of genes despite the constraint related to the possible number of probes, we performed preliminary experiments and compared four different designs; design 1: exons and introns were homogeneously covered by evenly distributed probes, with tiling space of 24 bp; design 2: only exons were covered by probes, and each exon was covered by 30–40 probes; design 3: exons and flanking intronic sequences were covered by 30–40 probes, and introns were covered by probes with spacing of 600 bp; design 4: exons and flanking intronic sequences were covered by 30–40 probes, and introns were covered by probes with a spacing of 100 bp. These experiments showed that design 2 was not suitable because of deviations from the expected baseline of $\log_2$ ratio corresponding to several consecutive probes. Design 3 and design 4 that require a much less number of probes than design 1 were appropriate to efficiently detect exonic CNMs. It appeared through these experiments that inclusion of a set of backbone probes covering intergenic regions would also contribute to overcome baseline irregularity.

We selected 26 genes routinely analyzed in our laboratory involved in heterogeneous groups of disorders and prone or suspected to be prone to CNMs (Tables 1 and 2). Genomic size and organization are highly variable between these genes, from the small tubulin genes with their four exons spanning about 4 kb of genomic DNA to the *DMD* gene spanning on 2.2 Mb and composed of 79 exons (Table 2). We

Table 1 Diseases selected and indications for CGH array analysis

Disease	Inheritance	Gene	Reported frequency of CNMs	Reference	Indications for CGH analysis
Myopathies					
Dystrophinopathies	XR	*DMD*	70%	Leiden muscular dystrophy pages (http://www.dmd.nl/)[2,3]	Diagnosis: first molecular screening (all index cases) Determination of boundaries
Sarcoglycanopathies	AR	*SGCA (alpha-SG)*	Rare (few cases)	Leiden muscular dystrophy pages (http://www.dmd.nl/)[17]	Patients with only one point mutation identified, or abnormal immunolabeling and no point mutation detected
	AR	*SGCB (beta-SG)*	Very rare (few cases)	Leiden muscular dystrophy pages (http://www.dmd.nl/)	Patients with only one point mutation identified, or abnormal immunolabeling and no point mutation detected
	AR	*SGCG (gamma-SG)*	Rare (few cases)	Leiden muscular dystrophy pages (http://www.dmd.nl/)[16,17]	Patients with only one point mutation identified, or abnormal immunolabeling and no point mutation detected
	AR	*SGCD (delta-SG)*	No reported case		Patients with only one point mutation identified, or abnormal immunolabeling and no point mutation detected
Emery Dreifuss syndrome	XR	*EMD*	Few cases	Leiden muscular dystrophy pages (http://www.dmd.nl/)[18]	Typical clinical syndrome, no point mutation in the *EMD* gene nor in *LAMA2* gene
Mental retardation					
Rett syndrome (RTT); Neonatal encephalopathy in males	XD	*MECP2*	5% of females with RTT 2% of males with severe encephalopathy Large CNMs	[20,21]	RTT females without point mutation in the *MECP2* gene. First molecular screening in males with severe encephalopathy
Rett variant with early epilepsy	XD	*CDKL5*	Rare, >20 cases	[22]	Atypical RTT Females without point mutation in the *MECP2* and *CDKL5* genes
RTT variant with congenital form	AD, *de novo*	*FOXG1*	Rare, <20 cases	[23,24]	Congenital variant of Rett syndrome without point mutations in the *MCEP2* and *FOXG1* genes
Rett-like syndrome	AD	*Netrin G1*	1 Case of translocation	[25]	Typical and atypical RTT patients without mutations in the *MECP2*, *CDKL5* and *FOXG1* genes
Rett-like syndrome	AD	*JNK3*	1 Case of translocation	[26]	Typical and atypical RTT patients without mutations in the *MECP2*, *CDKL5* and *FOXG1* genes
Fragile X syndrome	XD	*FMR1*	Rare (deletions)	[27]	Patients without expansion and with highly evocative phenotype
	XR	*FMR2*	Rare (deletions)	[28]	Patients without expansion and with highly evocative phenotype
Mental retardation because of *ARX*	XR	*ARX*	Rare (deletions)	[29]	Patients with no point mutation and with evocative phenotype
Lissencephalies and other cortical brain malformations	XD	*DCX*	Deletions and duplications described (large CNMs)	[30]	Patients with no point mutation and with evocative phenotype. Determination of boundaries
	XR	*OPHN1*	Rare		Patients with no point mutation and with evocative phenotype
	AD	*LIS1 (Pafah1b1)*	60% (Deletions) (large CNMs)	[30,31]	Patients with no point mutation and with evocative phenotype. Determination of boundaries
	AD	*TUBA1A*			Patients with no point mutation and with evocative phenotype
	AD	*TUBB2B*			Patients with no point mutation and with evocative phenotype
	AD	*TUBB3*			Patients with no point mutation and with evocative phenotype
	AD	*TUBB6*			Patients with no point mutation and with evocative phenotype
	AD	*TUBB5*			Patients with no point mutation and with evocative phenotype
Other diseases					
Cystic fibrosis or CFTR-related disorder	AR	*CFTR*	2.5–5%	Cystic fibrosis mutation database (www.genet.sickkids.on.ca/cftr/)[1,4]	Patients with cystic fibrosis or CFTR-related disorder heterozygous for a point mutation
Kallmann syndrome		*KAL1*	10% (large CNMs)	[32,33]	Males without point mutations in the 5 *KAL* genes Determination of boundaries
Hemophilia A	XR	*F8*	5–10%	[19]	Cases without recurrent intron 22 and intron 1 F8 inversions and without point mutations

Table 2 Human disease genes selected and design of the custom CGH array

Gene	Chromosome	Cytogenetic band	Gene location (UCSC Hg18-ref. seq gene)	Ref. seq	Gene size (kb)	Exons (N)	OMIM	Design (average tilling)
DMD	X	Xp21.2	31 047 266–33 267 647 (Dp427c)	NM_000109	2220	79	300 377	50 bp
SGCA (alpha-SG)	17	17q21	45 598 365–45 608 292	NM_000023.1	10	10	600 119	15 bp
SGCB (beta-SG)	4	4q12	52 581 618–52 599 242	NM_000232	17,6	6	600 900	15 bp
SGCG (gamma-SG)	13	13q12	22 653 060–22 797 304	NM_000231	144,2	8	608 896	40 bp exons, 300 bp introns
SGCD (delta-SG)	5	5q33-5q34	155 686 345–156 127 376	NM_000337	441	9	601 411	40 bp
EMD	X	Xq28	153 260 791–153 263 077	NM_000117	2,3	6	300 384	15 bp
MECP2	X	Xq28	152 940 458–153 016 382	NM_004992	76	4	300 005	15 bp
CDKL5	X	Xp22.13	18 353 646–18 581 670	NM_003159	228	21	300 203	40 bp exons, 300 bp introns
FOXG1	14	14q12	28 306 038–28 308 622	NM_005249.3	2,6	1	164 874	15 bp
NTNG1	1	1p13.3	107 484 152–107 825 998	NM_001113226.1	342	8	608 818	40 bp exons, 300 bp introns
JNK3 (MAPK10)	4	4q21.3	87 155 300–87 593 307	NM_138980.2	436,7	14	602 897	40 bp exons, 300 bp introns
FMR1	X	Xq27.3	146 801 201–146 840 333	NM_2024.4	39	17	300 624	40 bp exons, 300 bp introns
FMR2	X	Xq28	147 389 831–147 889 899	NM_2025	500	22	300 806	40 bp exons, 300 bp introns
ARX	X	Xq21.3	24 931 732–24 943 986	NM_007492	12	5	300 382	15 bp
DCX	X	Xq22.3q23	110 423 663–110 542 062	NM_178152	118	8	300 121	40 bp exons, 300 bp introns
OPHN1	X	Xq12	67 178 911–67 570 024	NM_002547	391	24	30 127	40 bp exons, 300 bp introns
LIS1 (Pafah1b1)	17	17p13.3	2 443 673–2 535 659	NM_000430	92	10	601 545	40 bp exons, 300 bp introns
TUBA1A	12	12p13.12	47 864 850–47 869 128	NM_006009	4	4	602 529	15 bp
TUBB2B	6	6p25	3 169 494–3 172 967	NM_178012	4	4	612 850	15 bp
TUBB3	16	16q24.3	88 517 188–88 530 006	NM_006086	4	4	602 661	15 bp
TUBB6	18	18p11.21	12 298 257–12 316 568	NM_032525	4	4	—	15 bp
TUBB5	6	6	—	NM_178014	4	4	—	15 bp
CFTR	7	7q31.2	116 907 253–117 095 954	NM_000492.3	189	27	602 421	20 bp
KAL1	X	Xp22.3	8 456 915–8 660 227	NM_000216	203,3	14	308 700	40 bp exons, 300 bp introns
F8	X	Xq28	153 717 258–153 904 192 (isoform a precursor)	NM_000132	187	26	306 700	40 bp exons, 300 bp introns
SHOX	X	Xp22.33	505 079–527 558	NM_000451.3	22	5	312 865	15 bp

50 bp: average tilling 50 bp in exons and introns.
15 bp: average tilling 15 bp in exons and introns.
20 bp: average tilling 20 bp in exons and introns.
40 bp exons, 300 bp introns: average tilling 40 bp in exons, 300 bp in introns.

developed a custom-targeted oligonucleotide-based CGH array containing 135 000 oligonucleotide probes covering genomic regions of these 26 genes, including their promoters and 2-kb upstream and downstream regions. A total of 344 exons, 26 promoters and corresponding intronic regions were then covered (Table 2). For each gene, the design of the 60mer probes was determined according to the characteristics of the gene and the preliminary experiments, in order to optimize visualization on CNMs (Table 2). For the 12 smaller genes (4–9 kb), exons and introns were covered by a HD of probes (tiling of 15 bp). For large genes, we chose to cover introns with a lower probe density (one probe each 300 bp) than exons (one probe each 40 bp) to save space on the array, except for the *CFTR* and *DMD* genes. For these two genes, intronic coverage was as dense as exonic coverage

(tiling of 20 bp for *CFTR* and 50 bp for *DMD*) to improve determination of breakpoints. Backbone probes were added at a lower density in intergenic sequences to obtain a stable baseline.

Validation of the custom CGH array by analysis of control DNAs
The performance of the custom CGH array was evaluated by using 38 DNA samples from patients with known CNMs previously identified by other techniques (Table 3). They ranged from small CNMs of one unique exon to entire gene rearrangements. Control patients were hemizygous, homozygous or heterozygous, and three somatic mosaicism were included.

Through the use of strict quality criteria, particularly the mad1.dr value below 0.23, no false-negative result was observed. All the 38

Table 3 Exonic copy-number mutations (CNMs) in *DMD*, *SGs*, *CFTR*, *CDKL5*, *DCX*, *LIS1*, *KAL* and *F8* genes used to validate the custom CGH array

Gene	N°	Mutation type	Conventional name	Name	Conventional methods	Status	Size of the CNM determined by CGH array
DMD	1[a]	Duplication	DMDdup2	c.32-?_93 + ?dup	QF-PCR, Q-PCR	He	250 kb
	2[a]	Duplication	DMDdup2_34	c.32-?_4845 + ?dup	QF-PCR	He	772 kb
	3[a]	Duplication	DMDdup3_9	c.94-?_960 + ?dup	QF-PCR	He	310.8 kb
	4[a]	Duplication	DMDdup17	c.1993-?2168 + ?dup	QF-PCR	He	11.1 kb
	5[a]	Duplication	DMDdup5_7	c.265-?_649 + ?del	QF-PCR	He	130 kb
	6[a]	Deletion	DMDdele45	c.6439-?_6614 + ?del	QF-PCR	He	151 kb
	7[a]	Deletion	DMDdele45	c.6439-?_6614 + ?del	QF-PCR	He	17.4 kb
	8[a]	Duplication	DMDdup48_49	c.6913-?_7200 + ?dup	QF-PCR	He	57 kb
	9[a]	Deletion	DMDdele51	c.7310-?_7542 + ?del	QF-PCR	He	49.8 kb
	10[a]	Deletion (mosaic)	DMDdele49_52	c.7099?_7660 + ?del	QF-PCR, RT-PCR	He	106 kb
	11[a]	Duplication	DMDdup55	c.8028-?_8217 + ?dup	QF-PCR	Ht	56 kb
	12[a]	Deletion	DMDdele46_47	c.6615-?_6912 + ?del	QF-PCR	He	23.9 kb
	13	Deletion	DMDdele53	c.7661-?_c.7872 + ?del	QF-PCR	He	18 kb
	14	Duplication	DMDdup7	c.531-?_649 + ?dup	QF-PCR	He	2.2 kb
	15[a]	Deletion	DMDdele45_47	c.6439-?_6912 + ?del	QF-PCR	Ht	149.4 kb
	16	Duplication (mosaic)	DMDdup61-62	c.9085-?_9224 + ?dup	QF-PCR, Q-PCR	Ht	2.3 kb
SGCG	17[a]	Deletion	SGdele5_6	c.386-?_578 + ? del/c.386-?_578 + ?del	QF-PCR	Ho	37.2 kb
	18[a]	Deletion	SGdele3	c.196-?_297 + ?del/c.196-?_297 + ?del	QF-PCR	Ho	16.2 kb
	19[a]	Deletion	SGdele7	c.579-?_702 + ?del/c.579-?_702 + ?del	QF-PCR	Ho	4.2 kb
	20[a]	Deletion	SGdele7	c.579-?_702 + ?del	QF-PCR	Ht	4.2 kb
SGCA	21[a]	Deletion	ASGdele7_8	c.748-?_983 + ?del	QF-PCR	Ht	1.2 kb
CFTR	22[a]	Deletion	CFTRdele3_10,14b_16	c.165-?_1584 + ?del;c.2620-?_2988 + ?del	MLPA	Ht	84 kb
	23[a]	Deletion	CFTRdele17a_17b	c.2989-977_3367 + 248del2515	MLPA	Ht	2.5 kb
	24[a]	Deletion	CFTRdele17a_18	c.2989-449_3468 + 644del5288	MLPA	Ht	5.2 kb
	25[a]	Deletion	CFTRdele2_3	c.54-?_273 + ?del	MLPA	Ht	21 kb
	26[a]	Deletion	CFTRdele22_23	c.3964-78_4242 + 577del1532	MLPA	Ht	1.5 kb
CDKL5	27	Deletion	CDKL5dele1	c.1-?_345 + ?del	MLPA	Ht	294. 4 kb
DCX	28	Deletion (mosaic 40:60, mutant:WT)	DCXdele4	c.706_5550_808 + 39del	QF-PCR, Q-PCR	He	6.2 kb
	29	Duplication	DCXdup4_7	c.705 + 18032_backbone	QF-PCR	He	81.6 kb
	30	Duplication	DCXdele2 Turner mosaique 20%	c.backbone_364 + 1054del	QF-PCR	Ht	7.9 kb
LIS1 (Pafah1b1)	31	Deletion Mieller Diecker	LIS1entire gene deletion	c.1-?_1233 + ?del	QF-PCR	Ht	740 kb
	32	Deletion	LIS1entire gene deletion AFH1B1dele4_11	c.1-?_1233 + ?del	QF-PCR	Ht	65.2 kb
	33	Deletion	LIS1entire gene deletion	c.1-?_1233 + ?del	QF-PCR	Ht	8.7 kb
	34	Deletion	LIS1entire gene deletion	c.1-?_1233 + ?del	QF-PCR	Ht	51.8 kb
KAL1	35	Deletion	*KAL1* entire deletion	c.1-?_2043 + ?del	Southern blot	He	2297 kb
	36		*KAL1* entire deletion	c.1-?_2043 + ?del	Southern blot	He	1600 kb
	37		Kal 1 dele3_13	c.256-?_1984 + ?del	Southern blot	He	91 kb
F8	38	Deletion	F8dele2_6	c.143-?_670 + ?del (new) c.86-?_613 + ?del (old)	PCR	He	38.6 kb

Abbreviations: N°, number; He, hemizygous patients; Ht, heterozygous individuals; Ho, homozygous patients.
A total of 38 control DNAs were studied in a blind trial. Nomenclature corresponds to the approved nomenclature. For hemophilia there are two nomenclatures: one that takes into account the peptide signal (the new nomenclature), one that does not take into consideration the peptide signal (the old nomenclature). Both are still used, to date, for diagnostic reports, and are indicated in the table.
[a]Control samples tested with the previous CGH array design.[13]

CNMs were accurately identified and precisely characterized, although rearrangements are of different size and scattered in different genomic regions (Table 3). The smallest deletion was a heterozygous 1.5 kb deletion of exons 22 and 23 within the *CFTR* gene that span 189 kb of genomic DNA (patient 26). We tested several monoexonic CNMs in the *DMD* gene, including a heterozygous duplication. All were correctly identified, as well as the frequent small CNV (1.4 kb) in intron 2 of this gene (Figure 1a). We also analyzed one DNA sample from chorionic villosities with a deletion in the *DMD* gene, which was correctly detected. A heterozygous deletion of exon 1 of the *CDKL5* gene (patient 27), which was difficult to detect by using MLPA because it appeared as a dosage reduction of a single peak among 45 signals, was easily visualized using CGH array, as it extends in fact on a large distal genomic region of 294 kb (Figure 2).

To assess the threshold of CNM detection with this procedure, we analyzed DNAs from three controls with somatic mosaicism. We reanalyzed DNA from a BMD patient with mosaicism in the *DMD* gene previously tested with the 4-plex design (patient 10, exons 49–52 deletion, 80:20 mutant:wild-type (WT)),[13] this mosaic deletion was correctly detected (data not shown). We also analyzed a male patient with lissencephaly who had a somatic mosaicism for a deletion in the *DCX* gene (patient 28; 40:60, mutant:WT). The deletion was detected with a log$_2$ ratio of −0.4 (data not shown). To refine our evaluation, we then tested DNA from a female carrier with a heterozygous duplication within the *DMD* gene in approximately 50% of cells (sample 16). CGH array analysis identified clearly the mosaicism as a gain of signal with a log$_2$ ratio of +0.3, instead of the 0.4 expected value for a heterozygous female (Figure 3).

Finally, relevance of inclusion of backbone probes was evaluated through analysis of deletions encompassing entire genes and extending outside of the gene regions. We tested four DNA samples from patients with lissencephaly because of heterozygous deletions of the whole *LIS1* gene (patients 31–34). The deletions were clearly identified, and boundaries roughly determined because of the low density of backbone probes in intergenic regions. The larger deletion was identified in a patient with Miller–Dieker syndrome, a contiguous gene disorder because of chromosome 17p13.3 deletion including *LIS1* (Figure 4a). We also analyzed two samples from patients with Kallmann syndrome because of large deletions of the entire *KAL1* gene (patients 35 and 36), including one with associated ocular albinism (patient 35). CGH array detected a large deletion of 2297 kb encompassing not only the whole *KAL1* gene but also neighboring genomic region including *GPR143* (Figure 4b), which is known to be associated, when mutated, with ocular albinism phenotypes.[34]

Use of the custom CGH array for diagnosis and genetic counseling
Following validation by using control DNAs with various gene CNMs, we decided to apply it as a diagnostic tool for patients with unknown molecular diagnosis and/or waiting for genetic counseling. CNM screening strategy varied according to genetic characteristics of the pathology (Table 1).

We applied the custom CGH array as a first diagnostic method to investigate the *DMD* gene in 262 DNA samples from patients (or relative women if DNA from patient is not available) for whom clinical, biological and histological data were compatible with the diagnosis of dystrophinopathy. We identified 102 deletions (90
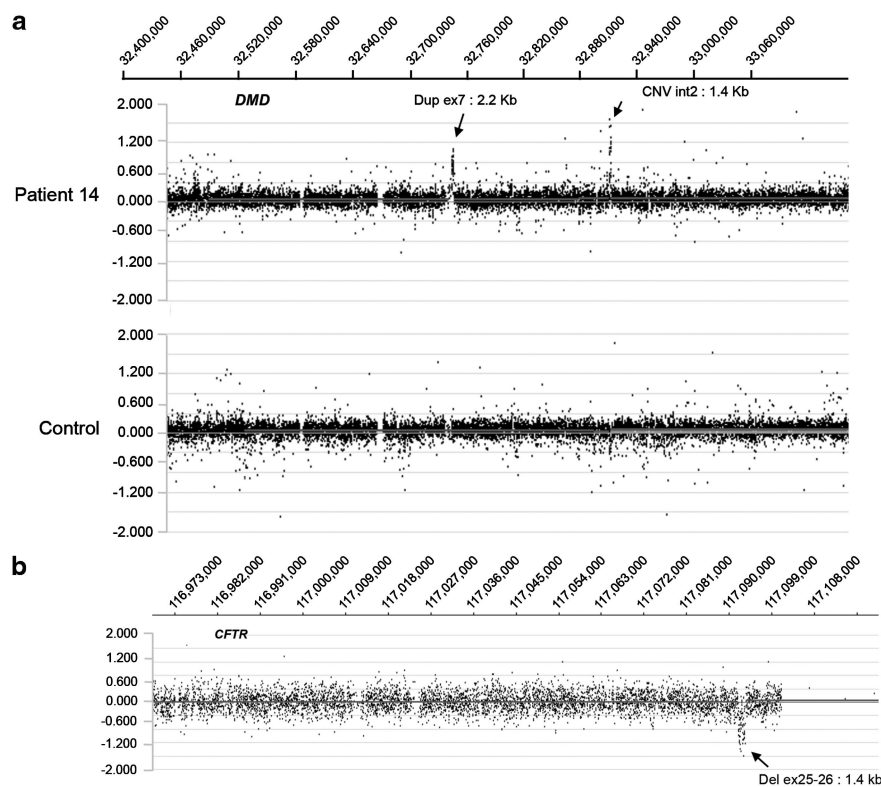
Figure 1 Detection by genomic CGH array of small CNMs and CNVs in different genes. (**a**) Detection of a hemizygous duplication of 2.2 kb encompassing exon 7 of the *DMD* gene and a duplication of 1.4 kb in intron 2 (patient 14). (**b**) Detection of a heterozygous 1.4 kb deletion carrying away exons 25 and 26 of the *CFTR* gene. The horizontal axis shows the position along the genome (NCBI36; Hg18) and the vertical axis the Cy3:Cy5 log$_2$ ratios. Patient sample was fluorescently labeled using Cy3 and control sample using Cy5. Control was sex matched with patient. The arrows indicate the location of the copy-number change.
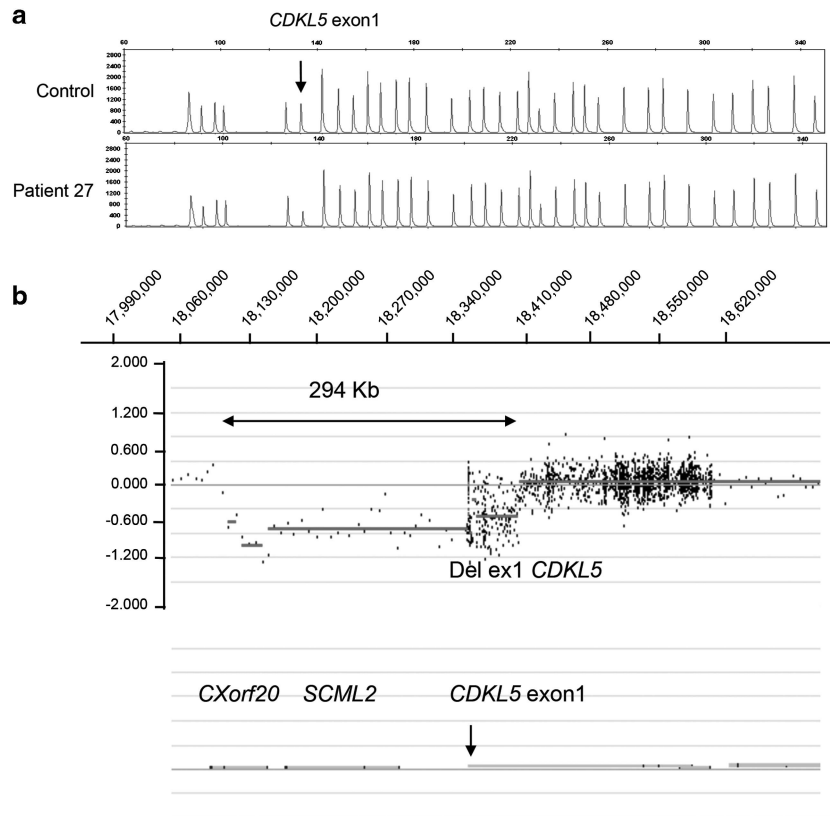
Figure 2 Heterozygous deletion of exon 1 of CDKL5 and neighboring genes (patient 27) is identified more reliably and accurately by CGH array compared with MLPA. (a) MLPA showing a reduction of the peak corresponding to CDKL5 exon 1 among CDKL5 exonic probe and control probe peaks. (b) Detection by genomic CGH array of a 294 kb deletion on chromosome X from intron 1 of the CDKL5 gene to the SCML2 and CXorf20 neighboring genes.

hemizygous and 12 heterozygous), 40 duplications (31 hemizygous and 9 heterozygous) and 2 complex rearrangements in the DMD gene. Single-exon CNMs were easily detected as they encompass in general part of flanking intronic regions. In contrast, rearrangements involving entire genes were visualized as a shift of the baseline corresponding to backbone probes in the intergenic regions. Interestingly, in a case of isolated duplication of exon 44 detected by cDNA analysis but not by QF-PCR, CGH array identified the duplication and showed that the proximal breakpoint was very close to exon 44. The primer used for QF-PCR was outside of the duplicated region, accounting for the false-negative result (data not shown). We detected, in two DMD families, a complex rearrangement, that is, CNMs involving different genomic parts of the gene and/or with more than two breakpoint junctions. The first case was the association of a duplication of exons 61–62, a duplication of exons 65–67 and abnormal values of $\log_2$ ratio for exons 68–79 probes suggesting a triplication. Exons 63–64 were normal. We confirmed by real-time Q-PCR that exon 63 was normal, exon 67 duplicated and exon 68 triplicated (Figure 5A). CGH array analyses performed in samples from three relative women of different generations showed that the rearrangement were stable and transmitted ad integrum. The other case of complex rearrangement in the DMD gene was identified in a heterozygous female and associated a duplication of exons 8–41, a duplication of exons 44–51 and a triplication of exons 42–43, confirmed by real-time Q-PCR (data not shown). Breakpoints characterized by CGH array were used to choose oligonucleotides

for sequencing 20 patients with various deletions in the DMD gene (article in preparation). The real breakpoints were distant in average of 190 bp for the proximal breakpoint (SD = 242 bp) and 267 bp for the distal breakpoint (SD = 275 bp). In most cases (n = 33 sequences), the difference was < 500 bp and sequencing was successful using the first set of primers chosen on the basis of CGH array boundaries (see Patients and methods section).

We also used the custom CGH array as a second-line diagnostic method to test DNA samples from 83 patients (or relatives when not available) with phenotypical data very suggestive of a disease and for whom only one point mutation or no point mutation was detected (18 Rett – typical or atypical – syndrome or neonatal encephalopathy, 42 CF or CFTR-related disorder, 23 severe hemophilia A). We identified 16 CNMs involving either MECP2 (n = 8), CDKL5 (n = 2), CFTR (n = 3) and the F8 gene (n = 3). CNMs involving MECP2 were two heterozygous deletions in Rett females, five hemizygous large duplications including the entire MECP2 gene and a large complex duplication–triplication (Figure 5B) in males with neonatal encephalopathy. Two heterozygous large deletions of CDKL5 and neighboring genomic regions were identified in females with atypical Rett syndrome. In CF patients, we identified three heterozygous CNMs involving CFTR, including the smallest CNM identified in this study (deletion of exons 25 and 26, 1.4 kb) (Figure 1b). CNMs were identified in the F8 gene in three families with severe hemophilia A, two heterozygous CNMs (one deletion of two exons, one duplication of eight exons) in carrier females (affected
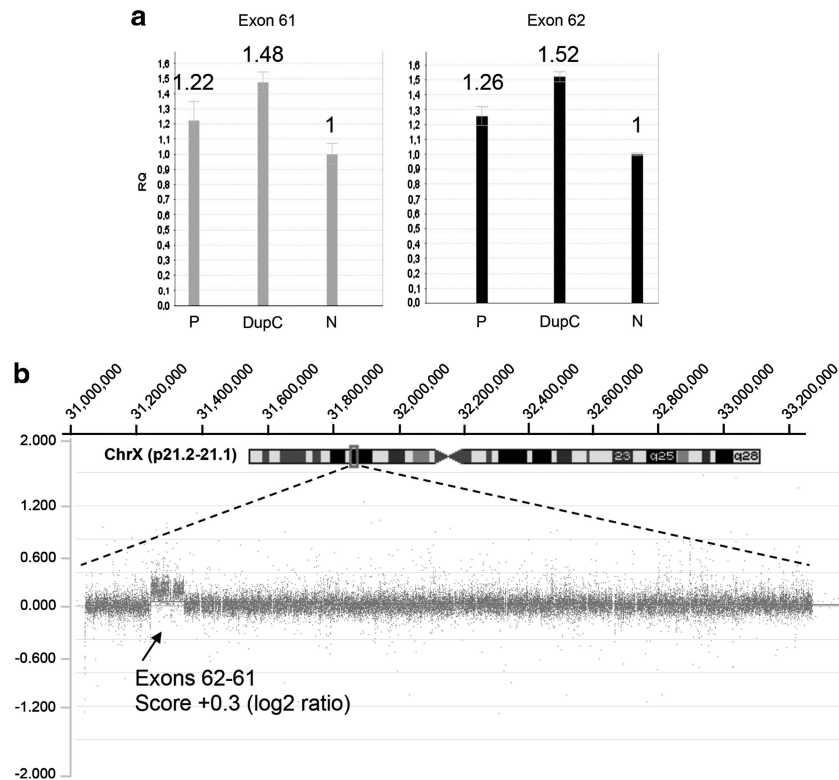
Figure 3 Detection of a somatic mosaicism corresponding to a duplication of exons 61 and 62 of the *DMD* gene in a carrier female (patient 16). (**a**) Real-time quantitative PCR of exons 61 and 62 performed in genomic DNA from blood of patient 16 (P), suggested that 50% of cells are heterozygous for the exons 61 and 62 duplication. DupC designates a control female heterozygous in all cells for duplication of exons 61 and 62 of the *DMD* gene. N indicates a normal control female. The vertical axis shows the relative quantification of the tested DNA compared with the normal control allele, in three different PCR experiments, the mean of the three quantifications is indicated above each peak. (**b**) Detection by genomic CGH array of the mosaic duplication. The arrow indicates the location of the duplication that extends on 2.3 kb encompassing exons 61 and 62. The $\log_2$ ratio is at $+0.3$, in accordance with the quantification of 50% of heterozygous mutant:wild-type cells.

males deceased) and a hemizygous deletion of five exons in a male patient.

We also used CGH array for genetic counseling in the family of patient 36 affected by Kallmann syndrome. Analysis of his unaffected mother revealed that the large deletion (whole *KAL1* gene and neighboring genomic region) is not present in leukocytes from the mother (Figure 4b).

DISCUSSION

CNM detection has improved considerably over the past 15 years. Initially, large rearrangements were characterized by Southern blot technology, which is a manual and time-consuming method. Appeared then fluorescent semi-quantitative PCR techniques and real-time quantitative PCR that are more accurate but do not allow simultaneous analysis of several genes. MAPH and MLPA techniques have initiated the change in the management of CNMs, allowing simultaneous analysis of various exons and genes responsible for similar diseases. Moreover, the emergence of commercial kits, providing better reproducibility, ensured the diffusion of MLPA approach. However, commercial MLPA kits do not allow concomitant analysis of large number of genes. The development of CGH array methodology constituted a real breakthrough. In this study, we describe implementation, validation and interests of a custom CGH array that analyses the 344 exons from 26 genes tested in our diagnostic laboratory.

For the array design, we decided to use variable density of probes depending on gene characteristics. For genes with small genomic size, we opted for a HD of probes in exons and introns. For the *DMD* and the *CFTR* genes, which are frequently analyzed, we chose a HD of probes in exons and introns to facilitate identification of intronic breakpoints. For the other large genes, we decided to place a lower density of probes in introns to save space on the chip. Intergenic regions were covered by backbone probes to maintain baseline stability between genes. This design appeared to us generally satisfactory in terms of sensitivity and specificity. By taking into account strict quality criteria, as the mad1dr value to validate each experiment, we did not have any false-negative results among the 38 control DNA samples. In a second phase of the study, we tested 345 cases not previously tested and identified 178 exonic CNMs in different genes (162 in *DMD*, 8 in *MECP2*, 2 in *CDKL5*, 3 in *CFTR* and 3 in the *F8* gene). Single-exon CNMs were clearly identified and sensitivity threshold was of 1.4 kb (but no smaller rearrangement was available in this series). In some genes like *KAL1*, CNMs are difficult to detect by PCR-based techniques because of a homologous gene (*KALP* on chromosome Y in the case of *KAL1*). We correctly identified three deletions of the *KAL1* gene, and detected one case of neomutation. The fact that oligonucleotide CGH array technique is based on the hybridization of DNA on multiple probes spread all along the target sequence allows to avoid false-negative and false-positive results obtained with the use of PCR techniques and MLPA. Our custom CGH array was efficient to determine the size of the
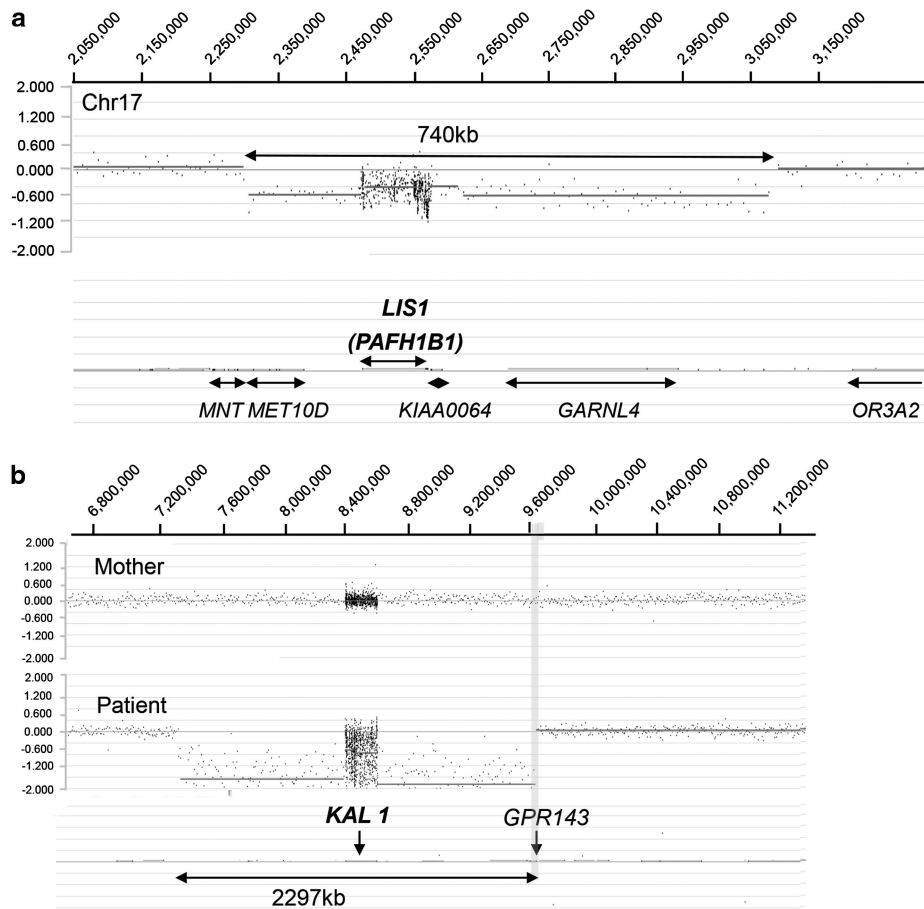
Figure 4 Detection of rearrangements involving whole genes and neighboring genomic regions. (**a**) Heterozygous deletion of 740 kb involving the entire *LIS1* (*PAFHB1*) gene and neighboring *MET1OD*, *KIAA0064* and *GARNL4* genes identified in patient 31 with a Mieller Diecker syndrome. The *MNT* and *OR3A2* genes are not deleted. (**b**) Hemizygous deletion of the entire *KAL1* gene in patient 35 with Kallman syndrome associated with ocular albinism, corresponding to a neomutation event. A 2297-kb deletion is detected in the patient sample and involves not only the *KAL1* gene, but also neighboring genomic region including the *GPR143* (*OA1*) gene. Mutations in this gene are known to be associated with ocular albinism phenotypes. The deletion is not detected in DNA sample from the mother.

deleted/duplicated fragments. Accuracy of this design was sufficient to allow targeted sequencing of breakpoints of CNMs in the *DMD* gene, as the average distance between breakpoints indicated by CGH and those determined by sequencing is < 300 bp. Large CNMs involving not only the gene of interest, but also neighboring genes, are not rare and identification of the other genes involved can be important for phenotype–genotype correlations. In our custom CGH array, presence of intergenic (backbone) probes allowed determining large rearrangements extending outside of the gene region as illustrated for *MECP2*, *LIS1* and *KAL1*. In a patient of our series with Kallmann syndrome associated with ocular albinism, CGH array allowed to identify a large deletion encompassing not only the *KAL1* gene but also the *GPR143* gene, which is known to be associated with ocular albinism phenotypes.[34]

Custom oligonucleotide CGH array technology also appears particularly efficient to detect and characterize precisely the nature of complex rearrangements that are rare but delicate to identify with current techniques including RT-PCR. In fact, we identified several cases of double duplications or complex rearrangements associating duplication and triplication of several exons in different genes (*DMD*, *CFTR* and *MECP2*), and demonstrated that their characterization is facilitated by simultaneous visualization of the entire rearrangement.

For instance, we identified in the *DMD* gene a particularly complex rearrangement that associates duplications and triplications of several, non-consecutive exons. To our knowledge, and while waiting for contributions from large-scale sequencing, only targeted CGH array approach allows to easily define extent and exons copy-numbers of complex rearrangements that is important for accurate familial studies (carrier determination and prenatal diagnosis) and for phenotype–genotype correlation studies. Somatic mosaicisms are even rarer, but their identification is important for genetic counseling. We tested a women suspected of having a somatic mosaicism of a duplication of two exons in the *DMD* gene. The clear visualization of the entire duplication in this woman allowed us to confirm that she was a carrier for the mutation and to provide a more appropriate genetic counseling. We also used CGH array for carrier status determination in some families without available index case. This approach was particularly conclusive in our experience for dystrophinopathies, hemophilia A and Kallmann syndrome carriers, but can be applied for all diseases with large rearrangements.

In addition to its efficiency and sensitivity, this custom CGH array is very powerful as it allows simultaneous analysis of a large number of exons (344 exons) corresponding to 26 disease genes. The 12-plex CGH array can detect CNMs, which are different in localization, type
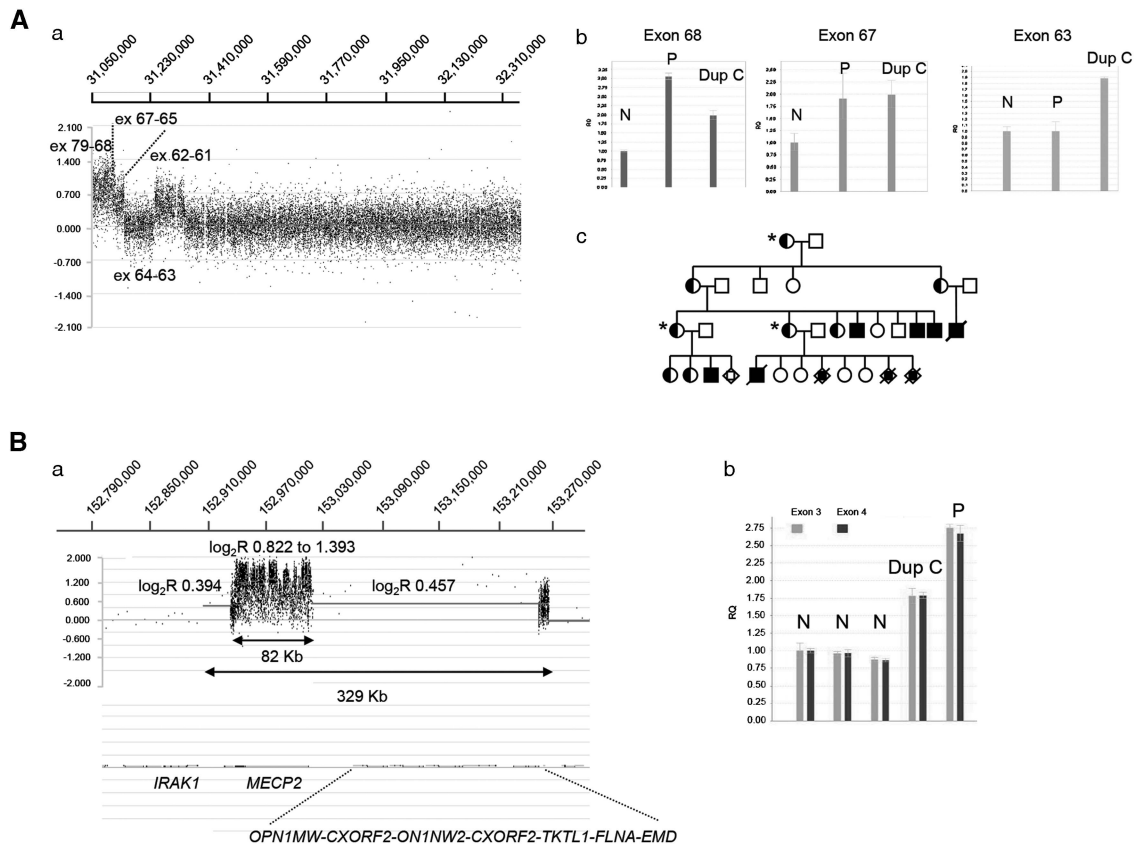
Figure 5 Identification of complex rearrangements. (**A**) Complex rearrangement in the DMD gene: duplication of exons 61–62, duplication of exons 65–67, triplication of exons 68–79. (**a**) CGH array. (**b**) Confirmation by real-time PCR of exons 63, 67 and 68. (**c**) Genealogical data. Tested women are indicated by an asterisk. The complex rearrangement was stably transmitted. (**B**) Triplication of the *MECP2* gene embedded within a duplication in a male patient with severe encephalopathy. (**a**) CGH array. The $\log_2$ ratios (indicated as $\log_2$ R) from 0.822 to 1.393 indicated a hemizygous triplication of a 82-kb region that includes the entire *MECP2* gene, embedded within a duplicated region of 329 kb involving several other genes. (**b**) Confirmation of the triplication by real-time PCR of exons 3 and 4 of the *MECP2* gene. Dup C, duplicated control; N, normal control; P, patient.

and size in a time <5 days for one experiment of 12 patients. If one wants to analyze the 26 genes by MLPA, which is a multistep approach, he would need to use eight different kits, with a working time of 2 days by kit and a cost at least double per sample. Moreover, some of the genes analyzed by our custom CGH array are not available on MLPA kits (in particular the *EMD* gene and tubulin genes), and in kits such as the *XLMR* genes, analyses are not exhaustive as probes are present for only some of the exons.

In conclusion, custom oligonucleotide-based CGH array has an undeniable input for diagnosis compared with conventional techniques by improving reliability and accuracy of CNM detection. The possibility of simultaneous analysis of several genes and its scalability make it a valuable tool for a new diagnostic approach of CNMs and should facilitate the molecular diagnosis of heterogeneous groups of diseases such as muscular dystrophies[16] or mental retardation. It is the first technology that allows determination of CNM boundaries precisely enough to guide targeted sequencing of breakpoints.[35] Possibility of high scale sequencing of breakpoints will bring real progress in understanding molecular mechanisms of rearrangements, searching for genotype–phenotype correlations and to guide certain therapeutic strategies.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

1 Quemener S, Chen JM, Chuzhanova N et al: Complete ascertainment of intragenic copy number mutations (CNMs) in the CFTR gene and its implications for CNM formation at other autosomal loci. Hum Mutat 2010; **31**: 421–428.
2 Deburgrave N, Daoud F, Llense S et al: Protein- and mRNA-based phenotype-genotype correlations in DMD/BMD with point mutations and molecular basis for BMD with nonsense and frameshift mutations in the DMD gene. Hum Mutat 2007; **28**: 183–195.
3 Tuffery-Giraud S, Beroud C, Leturcq F et al: Genotype-phenotype analysis in 2,405 patients with a dystrophinopathy using the UMD-DMD database: a model of nationwide knowledgebase. Hum Mutat 2009; **30**: 934–945.
4 Ferec C, Casals T, Chuzhanova N et al: Gross genomic rearrangements involving deletions in the CFTR gene: characterization of six new events from a large cohort of hitherto unidentified cystic fibrosis chromosomes and meta-analysis of the underlying mechanisms. Eur J Human Genet 2006; **14**: 567–576.
5 Casilli F, Di Rocco ZC, Gad S et al: Rapid detection of novel BRCA1 rearrangements in high-risk breast-ovarian cancer families using multiplex PCR of short fluorescent fragments. Hum Mutat 2002; **20**: 218–226.

6 Schouten JP, McElgunn CJ, Waaijer R, Zwijnenburg D, Diepvens F, Pals G: Relative quantification of 40 nucleic acid sequences by multiplex ligation-dependent probe amplification. *Nucleic Acids Res* 2002; **30**: e57.

7 White S, Kalf M, Liu Q *et al*: Comprehensive detection of genomic duplications and deletions in the DMD gene, by use of multiplex amplifiable probe hybridization. *Am J Hum Genet* 2002; **71**: 365–374.

8 del Gaudio D, Yang Y, Boggs BA *et al*: Molecular diagnosis of Duchenne/Becker muscular dystrophy: enhanced detection of dystrophin gene rearrangements by oligonucleotide array-comparative genomic hybridization. *Hum Mutat* 2008; **29**: 1100–1107.

9 Audrezet MP, Chen JM, Raguenes O *et al*: Genomic rearrangements in the CFTR gene: extensive allelic heterogeneity and diverse mutational mechanisms. *Hum Mutat* 2004; **23**: 343–357.

10 Hegde MR, Chin EL, Mulle JG, Okou DT, Warren ST, Zwick ME: Microarray-based mutation detection in the dystrophin gene. *Hum Mutat* 2008; **29**: 1091–1099.

11 Vissers LE, Veltman JA, van Kessel AG, Brunner HG: Identification of disease genes by whole genome CGH arrays. *Hum Mol Genet* 2005; **14**(Spec No. 2): R215–R223.

12 Bovolenta M, Neri M, Fini S *et al*: A novel custom high density-comparative genomic hybridization array detects common rearrangements as well as deep intronic mutations in dystrophinopathies. *BMC Genomics* 2008; **9**: 572.

13 Saillour Y, Cossee M, Leturcq F *et al*: Detection of exonic copy-number changes using a highly efficient oligonucleotide-based comparative genomic hybridization-array method. *Hum Mutat* 2008; **29**: 1083–1090.

14 Wong LJ, Dimmock D, Geraghty MT *et al*: Utility of oligonucleotide array-based comparative genomic hybridization for detection of target gene deletions. *Clin Chem* 2008; **54**: 1141–1148.

15 Staaf J, Torngren T, Rambech E *et al*: Detection and precise mapping of germline rearrangements in BRCA1, BRCA2, MSH2, and MLH1 using zoom-in array comparative genomic hybridization (aCGH). *Hum Mutat* 2008; **29**: 555–564.

16 Piluso G, Dionisi M, Del Vecchio Blanco F *et al*: Motor chip: a comparative genomic hybridization microarray for copy-number mutations in 245 neuromuscular disorders. *Clin Chem* 2011; **57**: 1584–1596.

17 Trabelsi M, Kavian N, Daoud F *et al*: Revised spectrum of mutations in sarcoglycanopathies. *Eur J Human Genet* 2008; **16**: 793–803.

18 Manilal S, Recan D, Sewry CA *et al*: Mutations in Emery-Dreifuss muscular dystrophy and their effects on emerin protein expression. *Hum Mol Genet* 1998; **7**: 855–864.

19 Vencesla A, Baena M, Garrido RP *et al*: F8 gene dosage defects in atypical patients with severe haemophilia A. *Haemophilia* 2012; **18**: 708–713.

20 Van Esch H, Bauters M, Ignatius J *et al*: Duplication of the MECP2 region is a frequent cause of severe mental retardation and progressive neurological symptoms in males. *Am J Hum Genet* 2005; **77**: 442–453.

21 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 Human Genet* 2007; **15**: 1218–1229.

22 Nemos C, Lambert L, Giuliano F *et al*: Mutational spectrum of CDKL5 in early-onset encephalopathies: a study of a large collection of French patients and review of the literature. *Clin Genet* 2009; **76**: 357–371.

23 Papa FT, Mencarelli MA, Caselli R *et al*: A 3 Mb deletion in 14q12 causes severe mental retardation, mild facial dysmorphisms and Rett-like features. *Am J Med Genet A* 2008; **146A**: 1994–1998.

24 Ariani F, Hayek G, Rondinella D *et al*: FOXG1 is responsible for the congenital variant of Rett syndrome. *Am J Hum Genet* 2008; **83**: 89–93.

25 Borg I, Freude K, Kubart S *et al*: Disruption of Netrin G1 by a balanced chromosome translocation in a girl with Rett syndrome. *Eur J Human Genet* 2005; **13**: 921–927.

26 Shoichet SA, Duprez L, Hagens O *et al*: Truncation of the CNS-expressed JNK3 in a patient with a severe developmental epileptic encephalopathy. *Hum Genet* 2006; **118**: 559–567.

27 Hammond LS, Macias MM, Tarleton JC, Shashidhar Pai G: Fragile X syndrome and deletions in FMR1: new case and review of the literature. *Am J Med Genet* 1997; **72**: 430–434.

28 Sahoo T, Theisen A, Marble M *et al*: Microdeletion of Xq28 involving the AFF2 (FMR2) gene in two unrelated males with developmental delay. *Am J Med Genet A* 2011; **155A**: 3110–3115.

29 Kato M, Das S, Petras K *et al*: Mutations of ARX are associated with striking pleiotropy and consistent genotype-phenotype correlation. *Hum Mutat* 2004; **23**: 147–159.

30 Haverfield EV, Whited AJ, Petras KS, Dobyns WB, Das S: Intragenic deletions and duplications of the LIS1 and DCX genes: a major disease-causing mechanism in lissencephaly and subcortical band heterotopia. *Eur J Human Genet* 2009; **17**: 911–918.

31 Lo Nigro C, Chong CS, Smith AC, Dobyns WB, Carrozzo R, Ledbetter DH: Point mutations and an intragenic deletion in LIS1, the lissencephaly causative gene in isolated lissencephaly sequence and Miller-Dieker syndrome. *Hum Mol Genet* 1997; **6**: 157–164.

32 Dode C, Hardelin JP: Kallmann syndrome. *Eur J Human Genet* 2009; **17**: 139–146.

33 Hardelin JP, Levilliers J, Young J *et al*: Xp22.3 deletions in isolated familial Kallmann's syndrome. *J Clin Endocrinol Metab* 1993; **76**: 827–831.

34 Bassi MT, Schiaffino MV, Renieri A *et al*: Cloning of the gene for ocular albinism type 1 from the distal short arm of the X chromosome. *Nat Genet* 1995; **10**: 13–19.

35 Ishmukhametova A, Van Kien PK, Mechin D *et al*: Comprehensive oligonucleotide array-comparative genomic hybridization analysis: new insights into the molecular pathology of the DMD gene. *Eur J Human Genet* 2012; **20**: 1096–1100.