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. 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. 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 mutations and are searched for in patients heterozygous for CFTR mutations and are searched for in patients heterozygous for CFTR mutations.

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 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. 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). 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. Finally, these scanning methods do not allow determination of CNM boundaries, and a limited number of

Keywords: CGH array, oligonucleotide probes; copy-number mutations; CNMs; complex rearrangements
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. 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 ours) 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.1,6

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 Tool. Any probe that did not map uniquely was removed except those in the pseudautosomal 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 previously13 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 InnoScan900 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, log2 ratio of the Cy3-labeled test sample versus Cy5-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 log2 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 log2 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
| Disease                                      | Inheritance | Gene   | Reported frequency of CNMs | Reference                                                                 | Indications for CGH analysis                                                                 |
|----------------------------------------------|-------------|--------|----------------------------|---------------------------------------------------------------------------|---------------------------------------------------------------------------------------------|
| Myopathies                                   |             |        |                            |                                                                           |                                                                                              |
| Dystrophinopathies                           | XR          | DMD    | 70%                        | Leiden muscular dystrophy pages [1],[2],[3]                               | Diagnosis: first molecular screening (all index cases) Determination of boundaries          |
| Sarcoglycanopathies                          | AR          | SGCA   | Rare (few cases)           | Leiden muscular dystrophy pages [1]                                       | Patients with only one point mutation identified, or abnormal immunolabeling and no point mutation detected |
| Sarcoglycanopathies                          | AR          | SGCB   | Very rare (few cases)      | Leiden muscular dystrophy pages [1]                                       | Patients with only one point mutation identified, or abnormal immunolabeling and no point mutation detected |
| Sarcoglycanopathies                          | AR          | SGCG   | Rare (few cases)           | Leiden muscular dystrophy pages [1]                                       | 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 [1]                                       | Typical clinical syndrome, no point mutation in the EMD gene nor in LAMA2 gene              |
| Mental retardation                           |             |        |                            |                                                                           |                                                                                              |
| Rett syndrome (RTT); Neonatal encephalopathy | XD          | MECP2  | 5% of females with RTT 2% of males with severe encephalopathy | [1],[21]                                                                  | Rett 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           |                                                                            | Atypical RTT Females without point mutation in the MECP2 and CDKL5 genes                  |
| Rett variant with congenital form            | AD, de novo | FOXG1  | Rare, < 20 cases           |                                                                            | Congenital variant of Rett syndrome without point mutations in the MECP2 and FOXG1 genes |
| Rett-like syndrome                           | AD          | Netrin G1 | 1 Case of translocation      |                                                                            | Typical and atypical RTT patients without mutations in the MECP2, CDKL5 and FOXG1 genes  |
| Rett-like syndrome                           | AD          | JNK3   | 1 Case of translocation     |                                                                            | Patients without expansion and with highly evocative phenotype                             |
| Fragile X syndrome                           | XD          | FMR1   | Rare (deletions)           |                                                                            | Patients without expansion and with highly evocative phenotype                             |
| Mental retardation because of ARX           | XR          | ARX    | Rare (deletions)           |                                                                            | Patients with no point mutation and with evocative phenotype                               |
| Lissencephalies and other cortical brain     | XD          | DCX    | Deletions and duplications described (large CNMs) |                                                                             | Patients with no point mutation and with evocative phenotype. Determination of boundaries |
| malformations                                 | XR          | OPHN1  | Rare                       |                                                                            |                                                                                              |
| AD                                           | LIS1 (Pafah1b1) | TUBA1A | 60% (Deletions) (large CNMs) | 30,31                                                                     | Patients with no point mutation and with evocative phenotype. Determination of boundaries |
| Other diseases                                |             |        |                            | Cystic fibrosis mutation database [1],[3]                                 | Patients with cystic fibrosis or CFTR-related disorder heterozygous for a point mutation |
| Cystic fibrosis or CFTR-related disorder     | AR          | CFTR   | 2.5–5%                     |                                                                            | Patients with cystic fibrosis or CFTR-related disorder heterozygous for a point mutation |
| Kallmann syndrome                            | XR          | KAL1   | 10% (large CNMs)           |                                                                            | Males without point mutations in the 5 KAL genes Determination of boundaries                |
| Hemophilia A                                 | XR          | FB     | 5–10%                      |                                                                            | Cases without recurrent intron 22 and intron 1 F8 inversions and without point mutations    |
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 \textit{CFTR} and \textit{DMD} genes. For these two genes, intronic coverage was as dense as exonic coverage (tiling of 20 bp for \textit{CFTR} and 50 bp for \textit{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

| Gene   | Chromosome | Cytogenetic band | Gene location (UCSC Hg18-ref. seq gene) | Ref. seq | Gene size (kb) | Exons (N) | OMIM | Design (average tiling) |
|--------|------------|------------------|----------------------------------------|----------|---------------|-----------|------|-------------------------|
| \textit{DMD} | X          | Xp21.2           | 31 047 266–33 267 647 (Dp427c)         | NM_000109 | 2220          | 79        | 300 377 | 50 bp                   |
| \textit{SGA} | 17         | 17q21            | 45 598 365–45 608 292                  | NM_000231 | 10            | 10        | 600 119 | 15 bp                   |
| \textit{SCG} | 4          | 4q12             | 52 581 618–52 599 242                  | NM_000232 | 17.6          | 6         | 600 900 | 15 bp                   |
| \textit{SCG} | 13         | 13q12            | 22 653 060–22 797 304                  | NM_000231 | 144.2         | 8         | 608 896 | 40 bp, 300 bp introns   |
| \textit{SCD} | 5          | 5q33-5q34        | 155 686 345–156 127 376                | NM_000337 | 441           | 9         | 601 411 | 40 bp                   |
| \textit{EMD} | X          | Xq28             | 153 260 791–153 263 077                | NM_000117 | 2.3           | 6         | 300 384 | 15 bp                   |
| \textit{MECP2} | X         | Xq28             | 152 940 458–153 016 382                | NM_004992 | 76            | 4         | 300 005 | 15 bp                   |
| \textit{CDK5} | X          | Xp22.13          | 18 353 646–18 581 670                  | NM_003159 | 228           | 21        | 300 203 | 40 bp, 300 bp introns   |
| \textit{FOXG1} | 14         | 14q12            | 28 306 038–28 308 622                  | NM_005249.3 | 2.6          | 1         | 164 874 | 15 bp                   |
| \textit{NTNG1} | 1          | 1p13.3           | 107 484 152–107 825 998                | NM_00111326.1 | 342       | 8         | 608 818 | 40 bp, 300 bp introns   |
| \textit{JNK3} | 4          | 4q21.3           | 87 155 300–87 593 307                  | NM_138980.2 | 436.7       | 14        | 602 897 | 40 bp, 300 bp introns   |
| \textit{FMR1} | X          | Xq27.3           | 146 801 201–146 840 333                | NM_2024.4  | 39            | 17        | 300 624 | 40 bp, 300 bp introns   |
| \textit{FMR2} | X          | Xq28             | 147 389 831–147 889 899                | NM_2025   | 500           | 22        | 300 006 | 40 bp, 300 bp introns   |
| \textit{AXR} | X           | Xq21.3           | 24 931 732–24 943 986                  | NM_007492 | 12            | 5         | 300 382 | 15 bp                   |
| \textit{DCX} | X           | Xq22.3q23        | 110 432 663–110 542 626                | NM_178152 | 118           | 8         | 300 121 | 40 bp, 300 bp introns   |
| \textit{OPHN1} | X          | Xq12             | 67 178 911–67 570 024                  | NM_002547 | 391           | 24        | 30 127  | 40 bp, 300 bp introns   |
| \textit{LIS1} | 17         | 17p13.3          | 2 443 673–2 535 659                    | NM_000430 | 92            | 10        | 601 545 | 40 bp, 300 bp introns   |
| \textit{TUBA1A} | 12         | 12p13.12         | 47 864 850–47 869 128                  | NM_006009 | 4             | 4         | 602 529 | 15 bp                   |
| \textit{TUBB2B} | 6          | 6p25             | 3 169 494–3 172 967                    | NM_178012 | 4             | 4         | 612 850 | 15 bp                   |
| \textit{TUBB3} | 16         | 16q24.3          | 88 517 188–88 530 006                  | NM_006086 | 4             | 4         | 602 661 | 15 bp                   |
| \textit{TUBB6} | 18         | 18p1.12          | 12 298 257–12 316 568                  | NM_032525 | 4             | 4         | —        | 15 bp                   |
| \textit{TUBB5} | 6           | 6               | —                                      | NM_178014 | 4             | 4         | —        | 15 bp                   |
| \textit{CFTR} | 7           | 7q31.2           | 116 907 253–117 095 954                | NM_000492.3 | 189       | 27        | 602 421 | 20 bp                   |
| \textit{KAL1} | X           | Xp22.3           | 8 456 915–8 660 227                    | NM_000216 | 203.3         | 14        | 308 700 | 40 bp, 300 bp introns   |
| \textit{SHOX} | X           | Xp22.3           | 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.

European Journal of Human Genetics
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\textsuperscript{1} | Mutation type | Conventional name | Name | Conventional methods | Status | Size of the CNM determined by CGH array |
|------|-------------------|----------------|------------------|------|----------------------|--------|----------------------------------------|
| DMD  | 1\textsuperscript{a} | Duplication    | DMDdup2          | c.32-?_93 + ?dup | QF-PCR, Q-PCR | He     | 250 kb                                 |
|      | 2\textsuperscript{a} | Duplication    | DMDdup2_34       | c.32-?_4845 + ?dup | QF-PCR | He     | 772 kb                                 |
|      | 3\textsuperscript{a} | Duplication    | DMDdup3_9        | c.94-?_960 + ?dup | QF-PCR | He     | 310.8 kb                               |
|      | 4\textsuperscript{a} | Duplication    | DMDdup17         | c.1993-?168 + ?dup | QF-PCR | He     | 11.1 kb                                 |
|      | 5\textsuperscript{a} | Duplication    | DMDdup5_7        | c.265-?_649 + ?del | QF-PCR | He     | 130 kb                                  |
|      | 6\textsuperscript{a} | Deletion       | DMDdele45        | c.6439-?_6614 + ?del | QF-PCR | He     | 151 kb                                  |
|      | 7\textsuperscript{a} | Deletion       | DMDdele45        | c.6439-?_6614 + ?del | QF-PCR | He     | 17.4 kb                                 |
|      | 8\textsuperscript{a} | Duplication    | DMDdup48_49      | c.691-?_7200 + ?dup | QF-PCR | He     | 57 kb                                   |
|      | 9\textsuperscript{a} | Deletion       | DMDdele51        | c.7310-?_7542 + ?del | QF-PCR | He     | 49.8 kb                                 |
|      | 10\textsuperscript{a} | Deletion (mosaic) | DMDdele49_52 | c.7099-?_7660 + ?del | QF-PCR, RT-PCR | He     | 106 kb                                 |
|      | 11\textsuperscript{a} | Duplication    | DMDdup55         | c.8028-?_8217 + ?dup | QF-PCR | Ht     | 56 kb                                   |
|      | 12\textsuperscript{b} | Deletion       | DMDdele46_47     | c.6615-?_6912 + ?del | QF-PCR | He     | 23.9 kb                                 |
|      | 13\textsuperscript{a} | Deletion       | DMDdele53        | c.7661-?_7872 + ?del | QF-PCR | He     | 18 kb                                   |
|      | 14\textsuperscript{a} | Duplication    | DMDdup7          | c.531-?_649 + ?dup | QF-PCR | He     | 2.2 kb                                  |
|      | 15\textsuperscript{a} | Deletion       | DMDdele45_47     | c.6439-?_6912 + ?del | QF-PCR, Ht | He     | 149.4 kb                               |
|      | 16\textsuperscript{a} | Duplication    | DMDdup61-62      | c.9085-?_9224 + ?dup | QF-PCR, Ht | He     | 2.3 kb                                  |
|      | 17\textsuperscript{a} | Deletion       | SGdele5_6        | c.386-?_578 + ?del | MLPA | Ho     | 37.2 kb                                 |
|      | 18\textsuperscript{a} | Deletion       | SGdele3          | c.196-?_297 + ?del | MLPA | Ho     | 16.2 kb                                 |
|      | 19\textsuperscript{a} | Deletion       | SGdele7          | c.579-?_702 + ?del | MLPA | Ho     | 4.2 kb                                  |
|      | 20\textsuperscript{a} | Deletion       | SGdele7          | c.579-?_702 + ?del | MLPA | Ho     | 4.2 kb                                  |
|      | 21\textsuperscript{a} | Deletion       | AGdele7_8        | c.748-?_983 + ?del | MLPA | Ho     | 1.2 kb                                  |
|      | 22\textsuperscript{a} | Deletion       | CFTRdele3_10_14b_16 | c.165-?_1584 + ?del | CFTRdele2_20_288 + ?del | MLPA | Ho     | 84 kb                                  |
|      | 23\textsuperscript{a} | Deletion       | CFTRdele17a_17b  | c.2989-?7367 + 248del2515 | MLPA | Ho     | 2.5 kb                                  |
|      | 24\textsuperscript{a} | Deletion       | CFTRdele17a_18   | c.2989-?4368 + 644del2528 | MLPA | Ho     | 5.2 kb                                  |
|      | 25\textsuperscript{a} | Deletion       | CFTRdele2_3      | c.54-?_273 + ?del | MLPA | Ho     | 21 kb                                   |
|      | 26\textsuperscript{a} | Deletion       | CFTRdele22_23    | c.3964-784242 + 577del1532 | MLPA | Ho     | 1.5 kb                                  |
| CDKL5| 27\textsuperscript{a} | Deletion       | CDKL5dele1       | c.1-?_345 + ?del | MLPA | Ht     | 294.4 kb                                |
| DCX  | 28\textsuperscript{a} | Deletion (mosaic:mutant:WT) | DCXdele4 | c.706-5550 + 808 + 39del | QF-PCR, Q-PCR | Ht     | 6.2 kb                                  |
|      | 29\textsuperscript{a} | Duplication    | DCXdup4_7        | c.705 + 18032 backbone | QF-PCR | Ht     | 81.6 kb                                 |
|      | 30\textsuperscript{a} | Duplication    | DCXdup2 Turner  | c.705 + 18032 backbone | QF-PCR | Ht     | 7.9 kb                                  |
|      | 31\textsuperscript{a} | Deletion       | LIS1entire gene deletion | c.1-?_1233 + ?del | QF-PCR | Ht     | 740 kb                                  |
|      | 32\textsuperscript{a} | Deletion       | LIS1entire gene deletion | c.1-?_1233 + ?del | QF-PCR | Ht     | 65.2 kb                                  |
|      | 33\textsuperscript{a} | Deletion       | LIS1entire gene deletion | c.1-?_1233 + ?del | QF-PCR | Ht     | 8.7 kb                                  |
|      | 34\textsuperscript{a} | Deletion       | LIS1entire gene deletion | c.1-?_1233 + ?del | QF-PCR | Ht     | 51.8 kb                                  |
| KAL  | 35\textsuperscript{a} | Deletion       | KALI entire deletion | c.1-?_2043 + ?del | Southern blot | He     | 2297 kb                                |
|      | 36\textsuperscript{a} | Deletion       | KALI entire deletion | c.1-?_2043 + ?del | Southern blot | He     | 1600 kb                                |
|      | 37\textsuperscript{a} | Deletion       | KALI_1 dele_3_13 | c.256-?_1984 + ?del | Southern blot | He     | 91 kb                                   |
| F8   | 38\textsuperscript{a} | Deletion       | F8dele2_6        | c.143-?_670 + ?del (new) | MLPA | He     | 38.6 kb                                 |

Abbreviations: N\textsuperscript{1}, 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.

\textsuperscript{a}Control samples tested with the previous CGH array design.\textsuperscript{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 villositys 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 log2 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 log2 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.

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 log2 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.
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...
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 \textit{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 \textit{DMD} and \textit{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 \textit{DMD}, 8 in \textit{MECP2}, 2 in \textit{CDKL5}, 3 in \textit{CFTR} and 3 in the \textit{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 \textit{KAL1}, CNMs are difficult to detect by PCR-based techniques because of a homologous gene (\textit{KALP} on chromosome Y in the case of \textit{KAL1}). We correctly identified three deletions of the \textit{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
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 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.

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 mosaicsisms 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 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 MET10D, 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.
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\(^\text{1}\) or mental retardation. It is the first technology that allows determination of CNM boundaries precisely enough to guide targeted sequencing of breakpoints.\(^\text{35}\)

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.

---

**ACKNOWLEDGEMENTS**

We are grateful for the patients and parents for their contribution in the study. We also thank the physicians for referring patients to our laboratory. We thank Sylvie Tuffery-Giraud for her help during the manuscript submission. This work was supported by the Assistance Publique–Hôpitaux de Paris, INSERM, CNRS, Université Paris Descartes and by the NMD-CHIP Consortium, a FP7 HEALTH project of the European Commission (Development of Targeted DNAChips for High-Throughput Diagnosis of Neuromuscular Disorders – Collaborative Project –FP7 Grant Agreement Number: HEALTH-F5-2008-223026).

---

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. Deburggrave N, Daoud F, Llerse 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, Leterouq 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.

---

**European Journal of Human Genetics**
6 Schouten JP, McGlougan 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 Audretz 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 Kong L, Dinnick D, Geragthy MT et al: Utility of oligonucleotide array-based comparative genomic hybridization for detection of target gene deletions. *Clin Chem* 2008; 54: 1141–1148.

15 Staal J, Tomgren T, Ramtech 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 Pituso 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 Chim Acta* 2011; 412: 1584–1596.

17 Trabeski M, Kavani N, Daoiu F et al: Revised spectrum of mutations in sarcoglycanopathies. *Hum J Pediatr Genet* 2008; 16: 793–803.

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

19 Vencesa A, Baena M, Garrido RP et al: FB 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 MECPR2 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 MECPR2 gene in Rett syndrome patients, including a familial case with a male proband. *Eur J Hum 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 dysmorphism and Rett-like features. *Am J Med Genet A* 2008, 146A: 1994–1998.

24 Ariani F, Hayek G, Roninella 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 Hum Genet* 2005; 13: 921–927.

26 Shiochet A, 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 Hum 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 Mol Genet* 2004; 23: 147–159.

30 Haverfield EV, Wilted 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 heterotopias. *Hum Mol Genet* 1997; 6: 157–164.

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 Hum 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* 1995; 76: 827–831.

34 Bassi MT, Schiaffino MV, Renieri A et al: Disruption of Netrin G1 by a balanced chromosome translocation in a girl with Rett syndrome. *Hum Genet* 2005; 118: 559–567.

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