

## ARTICLE

# Detection of small genomic imbalances using microarray-based multiplex amplifiable probe hybridization

Philippos C Patsalis<sup>\*1</sup>, Ludmila Kousoulidou<sup>1</sup>, Katrin Männik<sup>2</sup>, Carolina Sismani<sup>1</sup>, Olga Žilina<sup>2</sup>, Sven Parkel<sup>2</sup>, Helen Puusepp<sup>2</sup>, Neeme Tõnisson<sup>2</sup>, Priit Palta<sup>2</sup>, Maido Remm<sup>2</sup> and Ants Kurg<sup>2</sup>

<sup>1</sup>Department of Cytogenetics, The Cyprus Institute of Neurology and Genetics, Nicosia, Cyprus; <sup>2</sup>Department of Biotechnology, Institute of Molecular and Cell Biology, University of Tartu/Estonia Biocentre, Tartu, Estonia

Array-based genome-wide screening methods were recently introduced to clinical practice in order to detect small genomic imbalances that may cause severe genetic disorders. The continuous advancement of such methods plays an extremely important role in diagnostic genetics and medical genomics. We have modified and adapted the original multiplex amplifiable probe hybridization (MAPH) to a novel microarray format providing an important new diagnostic tool for detection of small size copy-number changes in any locus of human genome. Here, we describe the new array-MAPH diagnostic method and show proof of concept through fabrication, interrogation and validation of a human chromosome X-specific array. We have developed new bioinformatic tools and methodology for designing and producing amplifiable hybridization probes (200–600 bp) for array-MAPH. We designed 558 chromosome X-specific probes with median spacing 238 kb and 107 autosomal probes, which were spotted onto microarrays. DNA samples from normal individuals and patients with known and unknown chromosome X aberrations were analyzed for validation. Array-MAPH detected exactly the same deletions and duplications in blind studies, as well as other unknown small size deletions showing its accuracy and sensitivity. All results were confirmed by fluorescence *in situ* hybridization and probe-specific PCR. Array-MAPH is a new microarray-based diagnostic tool for the detection of small-scale copy-number changes in complex genomes, which may be useful for genotype–phenotype correlations, identification of new genes, studying genetic variation and provision of genetic services.

*European Journal of Human Genetics* (2007) 15, 162–172. doi:10.1038/sj.ejhg.5201738; published online 22 November 2006

**Keywords:** MAPH; microarray; submicroscopic aberrations; genomic analysis; chromosome X; bioinformatics

## Introduction

It is well known that various genetic imbalances can result in a wide range of pathological conditions such as mental

retardation, malformations, developmental delay, neurological disorders and congenital anomalies. Different methodologies have been introduced to clinical practice in order to detect these genetic imbalances and provide prenatal and postnatal diagnosis to numerous groups of patients by revealing the genetic cause and mechanism of their disease. Classical karyotyping allows whole-genome screening for chromosomal aberrations larger than 5 Mb, however, smaller aberrations are not detectable with the

\*Correspondence: Dr PC Patsalis, Department of Cytogenetics, The Cyprus Institute of Neurology and Genetics, PO Box 23462, 1683 Nicosia, Cyprus. Tel: +357 22 39 26 95; Fax: +357 22 39 27 93; E-mail: patsalis@cing.ac.cy

Received 6 July 2006; revised 28 September 2006; accepted 12 October 2006; published online 22 November 2006

existing G-banding analysis. Such small imbalances, called 'submicroscopic', require other methodologies that provide high-resolution analysis.

The development of fluorescence *in situ* hybridization (FISH) methods led to the detection of subtle cryptic chromosomal abnormalities smaller than 5 Mb.<sup>1,2,3</sup> Comparative genomic hybridization (CGH), using metaphase chromosomes permitted the detection of copy-number changes as small as 3 Mb.<sup>4</sup> The resolution of small size copy-number change detection was dramatically increased with the advancement of array-based CGH using BAC, PAC or cDNA clones,<sup>5-7</sup> achieving a resolution higher than 1 Mb.<sup>8</sup> To further increase, the sensitivity and specificity, the hybridization of genomic representations to CGH arrays has been developed.<sup>9</sup> An average resolution of 30 kb throughout the genome was achieved by arraying long oligonucleotide probes and hybridizing with representations from cancer and normal cells.<sup>10</sup> In addition, Affymetrix SNP arrays have recently been used for genome-wide detection of chromosomal regions with DNA copy-number alterations<sup>11,12</sup> and other commercial solutions exist with extreme resolution.<sup>13</sup> Currently, the tiling resolution BAC arrays covering the entire human genome with approximately one clone per 100 kb are also available.<sup>14</sup>

A novel DNA dosage detection method, called multiplex amplifiable probe hybridization (MAPH), enables accurate and reliable detection of changes in DNA copy-number with a theoretical resolution up to 100 bp within a defined region of interest.<sup>15</sup> The basic principle of the method is that the probes can be quantitatively recovered and amplified after hybridization on solid matrix. In practice, the studied genomic DNA is denatured, immobilized on a membrane and hybridized to a mixture of probes. Each probe is designed to be amplifiable and to target any unique coding or non-coding genomic sequence. As MAPH is fast, cost-effective and relatively easy to perform,<sup>16</sup> it has been used for the detection of small DNA copy-number changes.<sup>17</sup> A major drawback of MAPH is the fact that gel-based detection used in this method is only applicable to a limited number of probes. The majority of mentioned classical MAPH studies have included about 30–50 probes per assay.

We have applied classical MAPH to a novel microarray-based approach, aiming to contribute to the advancement in diagnosis of small copy-number changes throughout the human genome. The new array-MAPH method enables accurate and reliable determination of copy-number changes in virtually any targeted locus of human genome. It combines the flexibility, specificity and sensitivity of MAPH with the potential of high-resolution genomic analysis, provided by a microarray format. Here, we describe the detection of large and small genomic imbalances through the fabrication, interrogation and validation of a human chromosome X-specific array at the 238 kb median spacing of probes.

## Methods

Array-MAPH methodology is described in Figure 1. The methodology requires (a) probe selection and preparation, (b) microarray preparation and (c) array-MAPH hybridization and data analysis.

### Probe selection and preparation

Array-MAPH requires the use of specifically designed hybridization probes. Such hybridization probes should be: unique in the human genome; evenly spaced over the studied genomic region; similar in size (preferably 400–600 bp) and GC content (30–55%) to ensure similar hybridization conditions. In addition, these probes have to be amplified from the human genome, which requires design of PCR primers suitable for genomic PCR. To design locus-specific hybridization probes used in the current study, 1500 evenly spaced candidate regions from chromosome X were determined. The sequences (1000 bp) from each region were downloaded from the ENSEMBL database (NCBI Human Genome assembly build 35.1). In the first phase of the hybridization probe design, all candidate sequences were tested for their uniqueness by comparing them against each other with BLAST2 program.<sup>18</sup> Sequences that showed similarity score higher than 100 bits (~identical regions longer than 50 bp) with any other candidate region were rejected. Secondly, PCR primers for amplifying hybridization probes from the candidate sequences were designed with the modified Primer3 program<sup>19</sup> using the parameters shown in Supplementary Method 1. Thirdly, all designed PCR primers were tested for their uniqueness in the human genome. A substring consisting of 18 nucleotides from the each primer's 3' end was detached and the number of the occurrences of these substrings in the human chromosome sequences was counted as potential binding sites. If a pair of primers had in sum more than 10 putative binding sites or if both primers had more than three occurrences in the human genome, they were discarded together with corresponding probe region. Primers were also rejected, if a pair of primers was able to generate more than one PCR product from the human genome. Product was predicted if two primers bound onto opposite strands within 1000 bp from each other. In the last step, the remaining candidate probes were tested conclusively for their uniqueness in the human genome using the BLAST2 program. The DUST filter within BLAST2 algorithm was switched off with the - F F option, otherwise, some repeated regions in the human genome could remain unnoticed. Probes that showed similarity score higher than 120 bits (~identical regions longer than 60 bp) with regions other than their own correct binding site in the genome were removed.

From the designed candidate probes, 558 probes from the human chromosome X were finally selected on the basis of uniform spacing, maintaining a 150–350 kb



rapid amplification of each probe. Universal vector-derived primers were used to confirm cloning efficiency (Supplementary Method 2b). Clone identity was confirmed for every probe by PCR amplification using the probe-specific primer sets (Supplementary Method 2c). The successfully cloned candidate probes were cultured in LB medium with 40 µg/ml ampicillin and cryopreserved with 10% glycerol at -80°C. Array-MAPH amplifiable probes were developed by PCR amplification using PZA and PZB universal primers flanking the vector-cloning site (Supplementary Method 2b).

Array-MAPH target sequences were amplified for spotting onto the array by PCR using specific unique primers for every cloned probe (Supplementary Method 2c). The array-MAPH target sequences were then dissolved in 25% DMSO at final concentration of 30 ng/µl and spotted onto Genorama™ SAL-1 microarray slides (Asper Biotech Ltd, Tartu, Estonia) in duplicates with the use of VersArray ChipWriter Pro arrayer (BioRad Laboratories, Hercules, CA, USA). All probes were arrayed in random order, to minimize the possibility that a spatial artifact during array hybridization will be incorrectly interpreted as an aberration.

#### Array-MAPH hybridization and data analysis

Two micrograms of human test genomic DNA was immobilized on Hybond+ filters (GE Healthcare, Piscataway, NJ, USA) as described previously.<sup>15</sup> The filters were then hybridized with the array-MAPH amplifiable probes and washed to remove unbound and nonspecifically bound probes (Supplementary Method 3a). Each filter was placed into separate tube and all bound probes were recovered from filters by denaturing and quantitative PCR amplification (Supplementary Method 3b). The recovered probe mixture was purified and labeled by nick translation using aminoallyl-dUTP-s and later treatment with amino-reactive Cy3 dye (GE Healthcare) (Supplementary Method 3c). The recovered probe mixture was hybridized to the microarray for quantification using automated hybridization station HS-400 (Tecan Austria GmbH, Grödig/Salzburg, Austria) (Supplementary Method 3d). The microarrays were scanned using Affymetrix 428 microarray scanner (Affymetrix Inc., Santa Clara, CA, USA). Raw signal intensities were extracted with BaseCaller module of Genorama™ Genotyping Software 4.2 Package (Asper Biotech Ltd). As each target sequence was presented in duplicate, their average pixel fluorescence intensity was used for further analysis with the specifically designed software for array-MAPH, called MAPH-Stat, available on request for academic use. The same program was also used for resorting of probes into their genomic order.

Microarray signals were normalized using the MAPH-Stat program (between-slide normalization) with respect to the median of autosomal control probe-specific signals from the given microarray. In order to adjust the raw data to

normal distribution, which is required for correct calculation of confidence intervals (CIs), raw signal intensity values were converted to logarithmic scale by taking logarithm of signal intensity values with base 2. Then, average (AVERAGE) and CI 90% values were calculated for each (*i*) probe using data from the control panel containing signals from five DNA samples of cytogenetically controlled phenotypically normal individuals. This was carried out separately for male and female control panel. The use of CI 90% gives the theoretical false positive (two consecutive probes deviated from CI values to the same direction) rate of 0.5%.

The following transformation was made for signal intensity values from studied DNA calculated from the control panel:  $CONVERTED\_VALUE_i = VALUE_i / AVERAGE_i$ . CI values for each probe were converted by using the same formula. An example of data analysis is shown in the Supplementary Table 2.

Signal intensity values of at least two adjacent probes with the intensity above or below the CI 90% values were considered indicative for the potential copy number change in the analyzed region.

#### FISH and PCR confirmation analyses

FISH analyses were used to confirm duplication on chromosome X of patient A-2879 (Supplementary Figure 1) and PCR analyses were used to confirm the deletion in patients 220728, 22467 and A045 (Supplementary Method 4). For confirmation with FISH, BAC and PAC clones were obtained from the clone library of The Wellcome Trust Sanger Institute and cultured according to the provided instructions ([www.sanger.ac.uk](http://www.sanger.ac.uk)). *STS* gene-specific probe for FISH analysis was obtained from Vysis Inc. (Downers Grove, IL, USA). The protocol for BAC and PAC extraction was obtained from BACPAC Resources Center Home Page ([www.bacpac.chori.org](http://www.bacpac.chori.org)). Labeling and hybridization was performed using Nick Translation Kit (Vysis Inc.) according to the provided standard protocol. The confirmation with PCR analyses is described in Supplementary Method 4.

#### Results

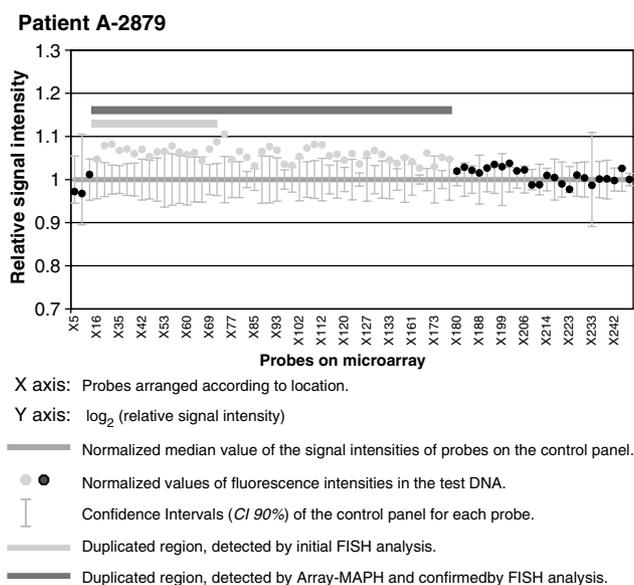
The first step towards the application of MAPH to a microarray format was the development of a chromosome X-specific array containing 558 probes spanning almost the entire chromosome X with median spacing of 238 kb and 107 control probes, representing all autosomal chromosomes. To simplify the *in silico* conducted probe design process and to make it reproducible, we prepared special software and tools for it. We also made a web interface called MAPHDesigner (<http://bioinfo.ebc.ee/MAPH>) for the developed programs. As web-based service, MAPHDesigner is freely available for all academic and non-academic users.

Probes were spotted onto microarray slides and used as hybridization targets. The chromosome X-specific

microarray was tested with the new array-MAPH methodology to analyze normal male and female DNA samples, as well as DNA samples from patients carrying known and unknown genomic imbalances. Each microarray with patient data is compared against a control panel of five normal individuals, which are analyzed on individual microarray. The signals from each array are normalized before comparison using the median signal intensity from the autosomal control probes from various experiments. Normalized signal values of the analyzed male and female samples were compared with CI 90% to the respective signal intensities for each of the array-MAPH probes from the control experiments.

Initial validation of array-MAPH was performed with normal DNA samples. All experiments were carried out in four replicates and gave comparable results. Replicate experiments are required to minimize possible biases related to experimental variations. DNA from a normal 46,XX control female subject was studied by comparing it against a control panel of five normal female subjects. As expected, the vast majority of the fluorescence intensities of chromosome X-specific probes remained within the calculated CI 90% values. From similar assay with four normal individuals, we estimated our empirical rate of false positives. This was carried out by counting all consecutive two probe deviations from CI 90% values trending into the same direction and was found to be ca. 1% on the average and 3% in the worst case. Secondly, DNA from a phenotypically normal male subject was compared to a control panel of five normal female subjects DNA samples. In this case, the normalized ratios for chromosome X-specific probes were clearly differentiated from autosomal as they remained below the CI 90%, demonstrating that the method reliably detects differences between one and two chromosomal copies.

To further validate the array-MAPH method, we used patient DNA samples with known chromosome X abnormalities. For various experiments, female patient DNA samples were compared to female control panels and male patient DNA samples to male control panels. Patient A-2879 is a female whom the initial cytogenetic analysis showed a chromosome Xp rearrangement. First FISH analyses using loci-specific probes for the *STS* and *SHOX* gene, the loci *DXZ1* and whole chromosome X paint revealed a complex aberration with a deletion of distal-most part of Xp22.3 and a proximal Xp22.31–p22.32 duplication. Array-MAPH was used to identify the size of the duplication. Our data showed the same duplication found by FISH; moreover, the duplicated area appeared to span a greater region than expected, extending to chromosomal bands Xp22.32–p22.12 (Figure 2). Second FISH analysis using the *STS* gene-specific probe, six BAC and one PAC clones confirmed the extended duplication identified by array-MAPH (Supplementary Figure 1). Our current probe list does not cover the distal-most part of Xp,



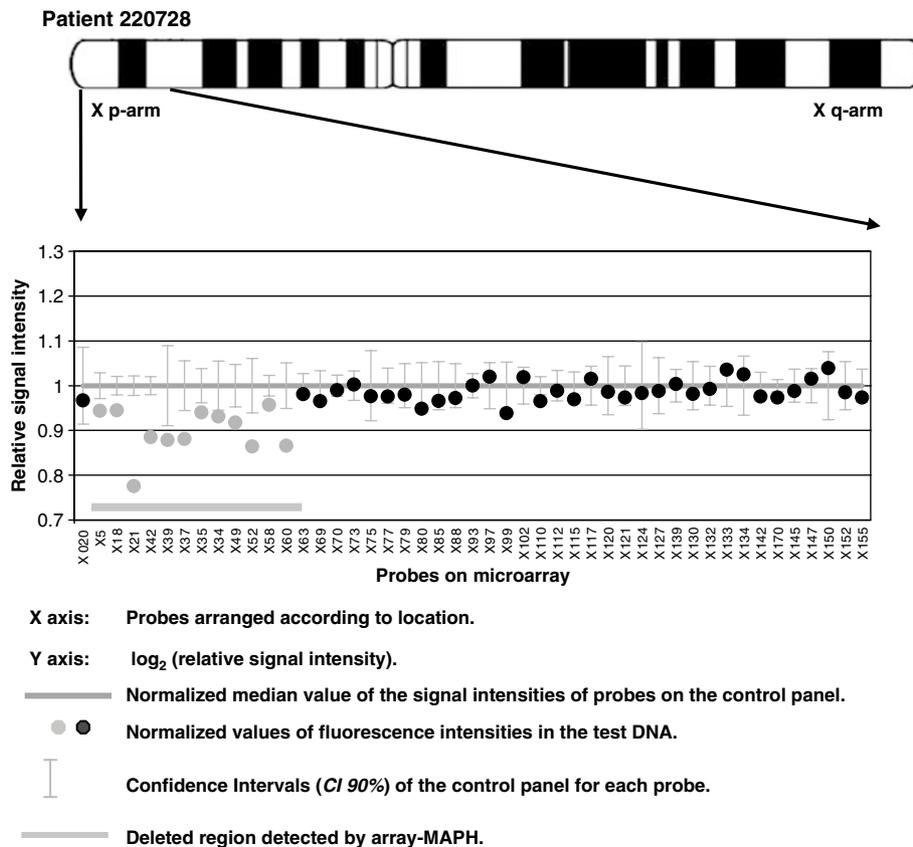
**Figure 2** Array-MAPH profile of patient A-2879.

so in this study, we could not analyze the small deleted region from initial FISH finding. Detailed description of the patient, clinical data and additional cytogenetic analysis will be reported elsewhere.

DNA samples of patients 220728 and 22467 were kindly provided by Dr J Vermeesch (Center for Human Genetics, Catholic University of Leuven, Belgium). The patient 220728 is a male carrying a deletion estimated initially to cover a region of maximum 7.5 Mb on the short arm of chromosome X detected by array-CGH and confirmed by FISH (Dr J Vermeesch, personal communication). The patient DNA was used in a blind test to validate the array-MAPH approach. Array-MAPH determined a 5.45–6 Mb Xp terminal deletion, spanning from probe X020 (2 771 275 bp) to X63 (8 443 336 bp) (Figure 3). Deletion breakpoints were confirmed by probe-specific and locus-specific PCR. The proximal breakpoint was estimated between 8 220 853 bp and 8 443 336 bp, whereas the distal breakpoint was between 2 445 055 bp and 2 771 275 bp (Figure 4). The discrepancy between array-CGH and array-MAPH results could be explained by the difference in the resolution of the two platforms.

Patient 22467 is a male, carrying a 1.5 Mb deletion on the short arm of chromosome X (Dr J Vermeesch personal communication). Array-MAPH showed a deletion between probes X34 (5 952 842 bp) and X58 (8 000 728 bp) (Figure 5), which was confirmed by probe-specific PCR (Figure 6).

Screening of 20 clinically affected male individuals from X-linked mental retardation families (unpublished data) resulted in the detection of a 500 kb deletion on Xq25 in patient A045 (Figure 7). The deletion was confirmed with probe-specific PCR (Figure 8a). Additional primers were



**Figure 3** Array-MAPH profile of patient 220728.

designed into the deleted region using the existing methodology for developing array-MAPH probes (Supplementary Table 3) and the breakpoints were refined by PCR (Figure 8b), estimating the deletion to be 23 kb in size spanning from 121 018 669 bp to 121 041 883 bp.

## Discussion

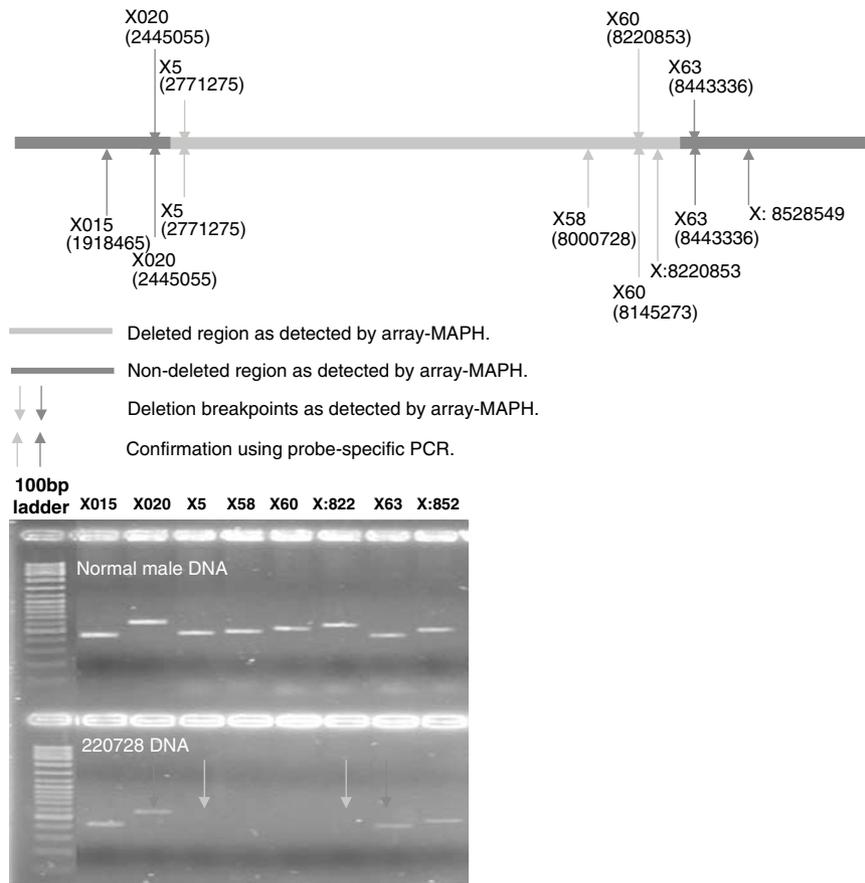
We describe the detection of large and small genomic imbalances using the new array-MAPH methodology. Array-MAPH was validated using a chromosome X-specific array with 558 probes, providing a median spacing of 238 kb over the entire human chromosome X. The 154-Mb human chromosome X was chosen because it is one of the best characterized chromosomes with complete sequence information available and provides an excellent model for studying X-linked genetic diseases. Initial validation of array-MAPH approach revealed certain limitations and advantages of the method.

Among the limitations of array-MAPH is the relatively large amount and concentration of starting DNA material needed. It was empirically established that best results can be obtained with 2  $\mu$ g of genomic DNA, whereas only 0.5–

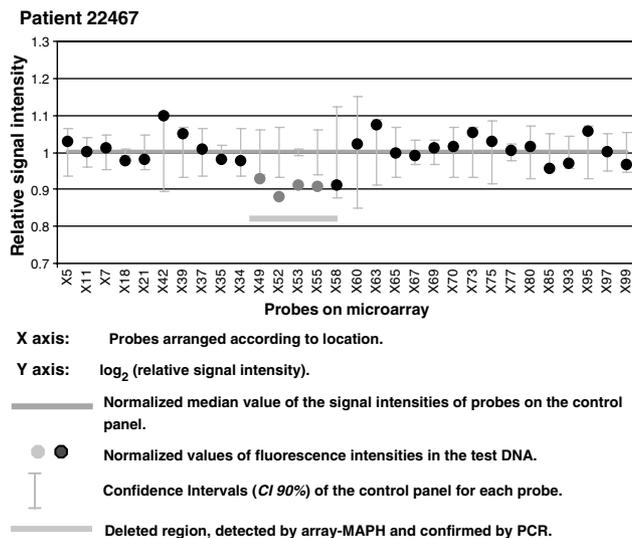
1  $\mu$ g is needed for gel-based MAPH,<sup>15</sup> 0.5  $\mu$ g for BAC/PAC array-CGH<sup>20</sup> and 250 ng for Affymetrix SNP arrays.<sup>11</sup> However, this amount is less than the 6  $\mu$ g needed for some oligonucleotide-based array platforms.<sup>21</sup> The necessary DNA quantity must be contained in a limited volume (less than 10  $\mu$ l) to enable spotting on small (2  $\times$  3 mm) filters. Therefore, DNA concentration should be at least 200 ng/ $\mu$ l. Nevertheless, one should consider that in case of array-MAPH the 2  $\mu$ g of DNA are required for preparing a filter and the same filter can be used for multiple analyses.

Array-MAPH involves two hybridization steps: the first one on filters and a second one on arrays. As a result, possible errors from both hybridizations are added to the final analysis. This problem is partially solved by uniform size and GC content of all probes, which allows maximum optimization of PCR and hybridization steps, avoiding errors and biases. In addition, the uniform size and GC content allows to overcome a potential hybridization issue of classical MAPH, where the probes have to be designed with different size for gel electrophoresis-based identification and quantification.

A potential drawback of single sample hybridization might be an increased influence of microarray-caused



**Figure 4** Probe-specific PCR confirmation of 5.67 Mb deletion for patient 220 728. Probes X5, X58, X60 and locus X:8220853, covering the region between 2771275 bp and 8220853 bp gave no product when amplified from patient's DNA, whereas probes X63 and locus X:8528549 proximal to the breakpoint and locus X015 and X020 distal to the breakpoint give normal PCR products.

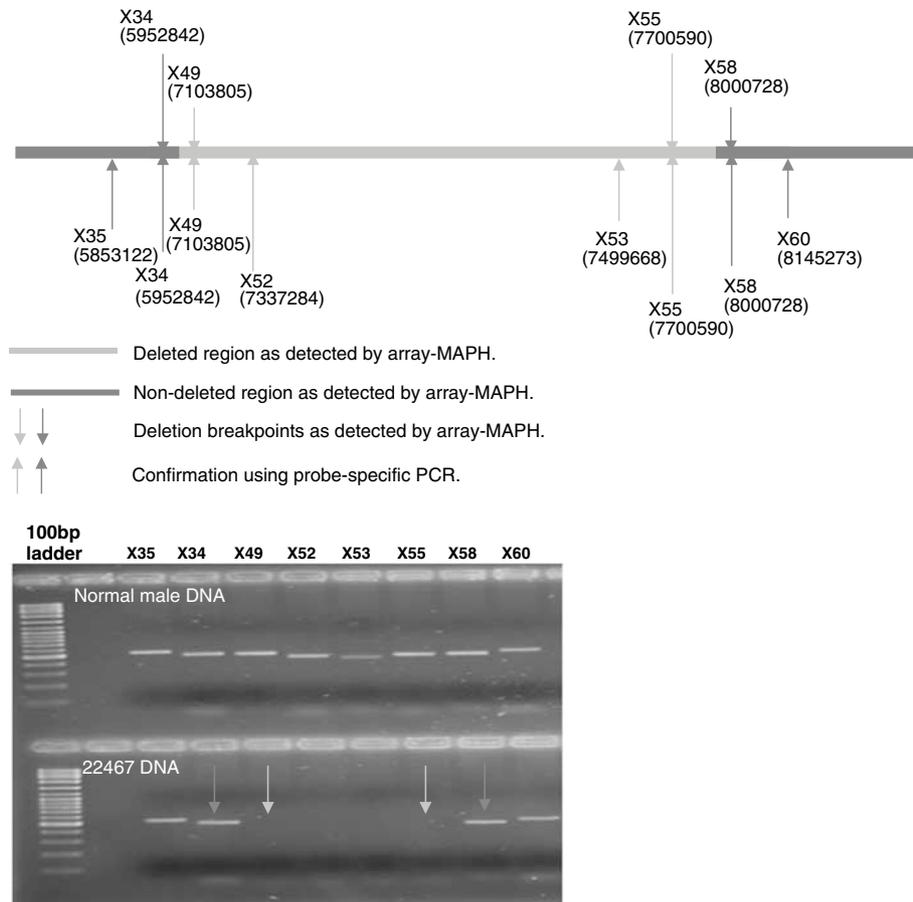


**Figure 5** Array-MAPH profile of patient 22467.

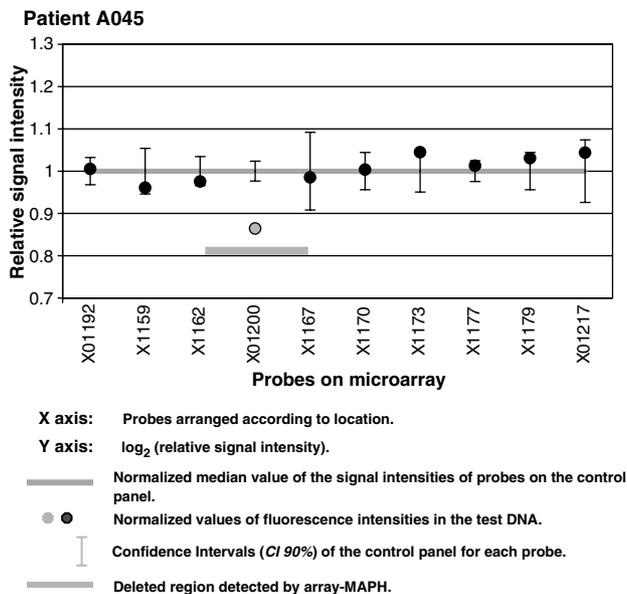
artefacts and variances. Although normalization and analysis is performed in a way that minimizes this type of experimental bias, still it is important to use automated hybridization systems to get as uniform hybridization efficiency as possible. At the same time, an automatic hybridization station will increase sample throughput.

A limiting factor for the clinical application of array-MAPH is that array-MAPH probes, similarly to oligonucleotides used for array-CGH, are extremely sensitive to DNA heterogeneity, which might cause an alteration of signal ratios. Therefore, one would expect that cancer cell line analysis might present certain challenges for array-MAPH, as tumor biopsies are often contaminated with normal tissue. This is also true for other samples with possible DNA contamination, such as amniotic fluid or fetal blood and mosaics.

A considerable advantage of the new approach is the nature of specifically designed probes it uses. *In silico*, probe development excludes repetitive or even partially

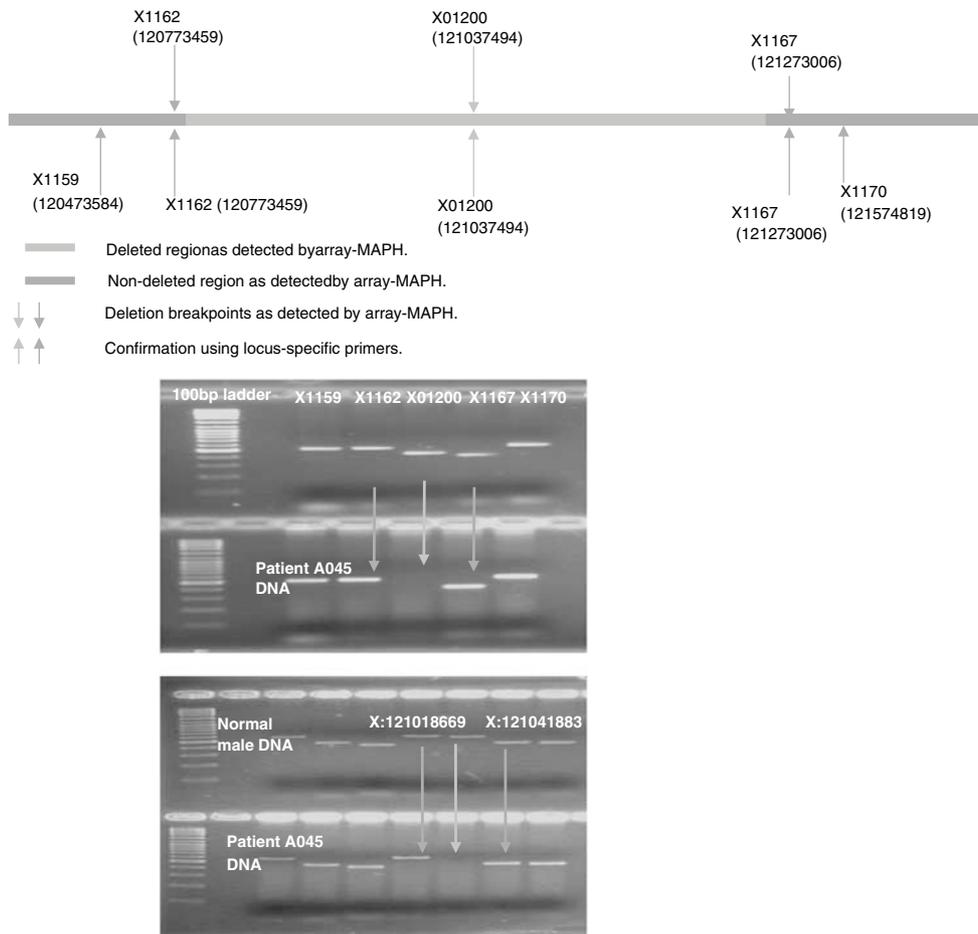


**Figure 6** Probe-specific PCR confirmation of 1.5 Mb deletion for patient 22467. Probes X49, X52, X53 and X55, covering the region between 7 103 805 bp and 7 700 590 bp give no product when amplified from patient's DNA, whereas probes X58 and X60 proximal to the breakpoint and probes X34 and X35 distal to the breakpoint give normal PCR products.



**Figure 7** Array-MAPH profile of patient A045.

redundant sequences and ensures specificity and sensitivity of array-MAPH, owing to 100% probe identity to the target sequence. Also, the uniqueness of probes enables to avoid the pseudoautosomal region on the chromosome X, which assures that probes used in such study will not result in the incorrect copy-number assessment for the corresponding chromosome X sequences. Similar probe development approach has been used also by other groups to design gene-specific strictly sequence defined microarrays for detection of diagnostically significant gene copy-number aberrations.<sup>22,23</sup> Comparison of array-CGH with genomic and PCR-based targets has shown that drawbacks of genomic clone-based approach caused mainly by the presence of redundant DNA sequences in the genomic clones can be circumvented by using short PCR-based microarray probes.<sup>24</sup> Probes for array-MAPH have the flexibility to target virtually any locus in the genome and cover large-scale copy-number variations (LCVs).<sup>25,26</sup> The relatively small size of probes (400–600 bp) allows detection of copy-number changes of small regions, facilitating the design of exon-specific and region-specific



**Figure 8** (a) Probe-specific PCR confirmation of 500 kb deletion for patient A045. Probe X01200 at 121 037 494 bp gives no product when amplified from patient's DNA, whereas probes X1159 and X1162 proximal to the breakpoint and probes X1167 and X1170 distal to the breakpoint give normal PCR products. (b) Amplification with additional primers narrows down the deletion to 23 kb between 121 018 669 bp and 121 041 883 bp. Probes located proximal to 121 018 669 bp and distal to 121 041 883 bp give normal PCR products, whereas probe located at 121 030 378 bp gives no product.

high-resolution arrays. This is also in a good agreement with data from others,<sup>23</sup> who have demonstrated that the shortest stretch of unique sequence, which allows reliable detection of single copy-number changes is 490 bp using fluorescently labeled total human DNA in the hybridization experiments.

Furthermore, one should consider the low complexity of DNA material hybridized to the microarray in case of array-MAPH and consequently increased concentration of probe complementary to target sequences. This is critical for obtaining reliable data,<sup>9,27,28</sup> as it significantly improves hybridization kinetics and increases signal to noise ratio. As genomic representations in case of ROMA technology,<sup>10</sup> array-MAPH also provides means to amplify the quantity of studied loci, thereby increasing signal to noise ratio during hybridization to arrayed probes.

Unlike array-CGH, where two different fluorophores are used, array-MAPH labeling involves a single fluorophore,

with reference signal values for comparison. Data from patient material is directly compared to CIs of the control data set. Recent large-scale studies show the presence of LCVs in several genomic regions of phenotypically normal individuals.<sup>25,26</sup> When searching for aberrations causing abnormal phenotypes, comparison of studied DNA against a panel composed of different pooled normal DNAs will enable more reliable analysis. However, LCVs can be investigated as well by limiting the heterogeneity of DNAs included in the control panel. In routine applications, a control panel will reduce the need for repeated analysis of normal DNAs, saving on materials and analysis time.

In contrast to clone-based technologies, the method does not rely on clone availability from BAC, PAC or other libraries, as probes can be rapidly and almost unrestrictedly selected from any location in genome. Having all probe-specific primers available, it is easy to carry out PCR-based confirmation procedures for any detected copy-number

alteration in studied DNA. Moreover, it is easy to confirm or reject candidate copy number changes and further refine the breakpoints by designing more primers using the existing software tools developed for MAPH probe design. Using this approach, we confirmed the deletion and its breakpoints for patient A045. Despite the fact that in general, single probe deviations as in case of A045, are considered to be 'false positive' and at present, we rely on two or more consecutive probes to warrant a copy-number change, there is always the option of simultaneous investigation of 10–20 loci, that showed 'false positive' ratios. Therefore, in some cases, even single probe alterations, classified as 'false positive' can be further studied and smaller deletions or duplications detected.

Our current data demonstrate the ability of array-MAPH followed by locus-specific PCR to detect deletions and duplications as small as 23 kb (patient A045). Nevertheless, the array-MAPH platform itself carries the potential of even higher resolution considering the 500 bp probe size and the flexibility to vary probe spacing according to the scientific question to be answered.

Array-MAPH can be useful for several applications in medical genomics. Even though array-MAPH can easily be applied to diagnosis of known genetic syndromes caused by copy-number alterations, it is a particularly useful tool for high-resolution screening of patients with abnormal phenotype and no genetic findings, detection of microdeletions, microduplications, genotype–phenotype correlations, gene identification and studying genetic variation. Further development and optimization of the technique may result in improved cost-effectiveness, reproducibility and dynamics, thereby revealing the full potential of array-MAPH platform.

### Acknowledgements

This work was funded by the grants 30/2001 from the Cyprus RPF, QLRT-2001-01810 from the EURO-MRX EU, 5467 from the Estonian Science Foundation, by 0182649s04 from the Estonian Ministry of Education and Research and by 070191/Z/03/Z from the Wellcome Trust International Senior Research Grant. We thank D Andreou, C Tryfonos, E Hadjiyanni, C Pitta, C Antoniadis, S Bashiardes and G Slavin for their contribution. Many thanks to Professor J Vermeesch, Dr K Öunap and Dr R Žordania for the provision of DNA samples and the Wellcome Trust Sanger Institute for the provision of BAC/PAC clones.

### References

- 1 Test and Technology Transfer Committee: Technical and clinical assessment of fluorescence *in situ* hybridization: an ACMG/ASHG position statement. I. Technical considerations. *Genet Med* 2000; **2**: 356–361.
- 2 Patsalis PC, Evangelidou P, Charalambous S, Sismani C: Fluorescence *in situ* hybridization characterization of apparently balanced translocation reveals cryptic complex chromosomal rearrangements with unexpected level of complexity. *Eur J Hum Genet* 2004; **12**: 647–653.
- 3 Knight SJ, Flint J: Perfect endings: a review of subtelomeric probes and their use in clinical diagnosis. *J Med Genet* 2000; **37**: 401–409.
- 4 Kallioniemi A, Kallioniemi OP, Sudar D *et al*: Comparative genomic hybridization for molecular cytogenetic analysis of solid tumors. *Science* 1992; **258**: 818–821.
- 5 Pinkel D, Seagraves R, Sudar D *et al*: High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays. *Nat Genet* 1998; **20**: 207–211.
- 6 Pollack JR, Perou CM, Alizadeh AA *et al*: Genome-wide analysis of DNA copy-number changes using cDNA microarrays. *Nat Genet* 1999; **23**: 41–46.
- 7 Solinas-Toldo S, Lampel S, Stilgenbauer S *et al*: Matrix-based comparative genomic hybridization: biochips to screen for genomic imbalances. *Genes Chromosomes Cancer* 1997; **20**: 399–407.
- 8 Snijders AM, Nowak N, Seagraves R *et al*: Assembly of microarrays for genome-wide measurement of DNA copy number. *Nat Genet* 2001; **29**: 263–264.
- 9 Lucito R, West J, Reiner A *et al*: Detecting gene copy number fluctuations in tumor cells by microarray analysis of genomic representations. *Genome Res* 2000; **10**: 1726–1736.
- 10 Lucito R, Healy J, Alexander J *et al*: Representational oligonucleotide microarray analysis: a high-resolution method to detect genome copy number variation. *Genome Res* 2003; **13**: 2291–2305.
- 11 Zhao X, Li C, Paez JG *et al*: An integrated view of copy number and allelic alterations in the cancer genome using single nucleotide polymorphism arrays. *Cancer Res* 2004; **64**: 3060–3071.
- 12 Huang J, Wei W, Zhang J *et al*: Whole genome DNA copy number changes identified by high density oligonucleotide arrays. *Hum Genomics* 2004; **1**: 287–299.
- 13 Selzer RR, Richmond TA, Pofahl NJ *et al*: Analysis of chromosome breakpoints in neuroblastoma at sub-kilobase resolution using fine-tiling oligonucleotide array CGH. *Genes Chromosomes Cancer* 2005; **44**: 305–319.
- 14 Ishkanian AS, Malloff CA, Watson SK *et al*: A tiling resolution DNA microarray with complete coverage of the human genome. *Nat Genet* 2004; **36**: 299–303.
- 15 Armour JA, Sismani C, Patsalis PC, Cross G: Measurement of locus copy number by hybridisation with amplifiable probes. *Nucleic Acids Res* 2000; **28**: 605–609.
- 16 Hollox EJ, Akrami SM, Armour JA: DNA copy number analysis by MAPH: molecular diagnostic applications. *Expert Rev Mol Diagn* 2002; **2**: 370–378.
- 17 Patsalis PC, Kousoulidou L, Sismani C, Mannik K, Kurg A: MAPH: from gels to microarrays. *Eur J Med Genet* 2005; **48**: 241–249.
- 18 Altschul SF, Madden TL, Schaffer AA *et al*: Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 1997; **25**: 3389–3402.
- 19 Rozen S, Skaletsky H: Primer3 on the WWW for general users and for biologist programmers. *Methods Mol Biol* 2000; **132**: 365–386.
- 20 Cai WW, Mao JH, Chow CW, Damani S, Balmain A, Bradley A: Genome-wide detection of chromosomal imbalances in tumors using BAC microarrays. *Nat Biotechnol* 2002; **20**: 393–396.
- 21 Barrett MT, Scheffer A, Ben-Dor A *et al*: Comparative genomic hybridization using oligonucleotide microarrays and total genomic DNA. *Proc Natl Acad Sci USA* 2004; **101**: 17765–17770.
- 22 Mantripragada KK, Buckley PG, Jarbo C, Menzel U, Dumanski JP: Development of NF2 gene specific, strictly sequence defined diagnostic microarray for deletion detection. *J Mol Med* 2003; **81**: 443–451.
- 23 Mantripragada KK, Thureson AC, Piotrowski A *et al*: Identification of novel deletion breakpoints bordered by segmental duplications in the NF1 locus using high resolution array-CGH. *J Med Genet* 2006; **43**: 28–38.
- 24 Mantripragada KK, Tapia-Paez I, Blennow E, Nilsson P, Wedell A, Dumanski JP: DNA copy-number analysis of the 22q11 deletion-syndrome region using array-CGH with genomic and PCR-based targets. *Int J Mol Med* 2004; **13**: 273–279.

- 25 Iafrate AJ, Feuk L, Rivera MN *et al*: Detection of large-scale variation in the human genome. *Nat Genet* 2004; **36**: 949–951.
- 26 Sebat J, Lakshmi B, Troge J *et al*: Large-scale copy number polymorphism in the human genome. *Science* 2004; **305**: 525–528.
- 27 Lucito R, Nakimura M, West JA *et al*: Genetic analysis using genomic representations. *Proc Natl Acad Sci USA* 1998; **95**: 4487–4492.
- 28 Kennedy GC, Matsuzaki H, Dong S *et al*: Large-scale genotyping of complex DNA. *Nat Biotechnol* 2003; **21**: 1233–1237.

Supplementary Information accompanies the paper on European Journal of Human Genetics website (<http://www.nature.com/ejhg>)