Rapid detection of genomic imbalances using micro-arrays consisting of pooled BACs covering all human chromosome arms

Jeroen Knijnenburg, Marja van der Burg, Philomeen Nilsson1, Hans Kristian Ploos van Amstel1, Hans Tanke and Károly Szuhai*

Department of Molecular Cell Biology, Leiden University Medical Center; Wassenaarseweg 72, 2333 AL Leiden, The Netherlands and 1Department of Medical Genetics, Utrecht Medical Center, Lundlaan 6, 3584 EA Utrecht, The Netherlands

Received August 12, 2005; Revised and Accepted September 27, 2005

ABSTRACT

A strategy is presented to select, pool and spot human BAC clones on an array in such a way that each spot contains five well performing BAC clones, covering one chromosome arm. A mini-array of 240 spots was prepared representing all human chromosome arms in a 5-fold as well as some controls, and used for comparative genomic hybridization (CGH) of 10 cell lines with aneusomies frequently found in clinical cytogenetics and oncology. Spot-to-spot variation within five replicates was below 6% and all expected abnormalities were detected 100% correctly. Sensitivity was such that replacing one BAC clone in a given spot of five by a BAC clone from another chromosome, thus resulting in a change in ratio of 20%, was reproducibly detected. Incubation time of the mini-array was varied and the fluorescently labelled target DNA was diluted. Typically, aneusomies could be detected using 30 ng of non-amplified random primed labelled DNA amounts in a 4 h hybridization reaction. Potential application of these mini-arrays for genomic profiling of disseminated tumour cells or of blastomeres for preimplantation genetic diagnosis, using specially designed DNA amplification methods, are discussed.

INTRODUCTION

Detection of gene amplifications and deletions by comparative genomic hybridization (CGH) is typically performed using arrays consisting of spotted BAC clones that adequately cover the genome of interest. In the human situation arrays of 3.5 k BAC elements are mostly used (1). Recently, arrays with ten times more BAC elements (35 k) have been successfully applied to detect aberrations at higher resolution (2). Since spotted arrays are relatively large (several cm²), a relatively high amount of fluorescently labelled target DNA is required to overcome diffusion barriers and achieve equilibrium conditions; prerequisites for a successful array-CGH experiment.

Consequently, hybridization times are in practice at least 24 h or longer, and more important enough assay material must be available to isolate sufficient amounts of DNA. In many cases however, the use of DNA amplification methods is indicated to generate sufficient copies of fluorescently labelled target DNA (typically 1–4 µg). Whereas methods such as DOP and ligation-mediated PCR have been fine-tuned such that even single cells can be assayed by conventional CGH on metaphase chromosomes (3,4), conditions for array-CGH are different. So far array-CGH of <100 cells has been very cumbersome, despite the use of a variety of successful DNA amplification strategies. One reason for this is the variation of the amplification process that leads to increasing differences in DNA representation with number of amplification cycles becomes apparent when the readout is performed at higher genomic resolution i.e. using BAC arrays with at least 1 MB resolution. In other words, presuming that the quality of the starting material is optimal, the variation of the mean signal ratio of all targets (spots) increases with increasing number of amplification steps and increasing array complexity, i.e. the number of DNA sequence elements spotted on the array. Although abnormalities can be detected in even relatively noisy data at the cost of the actual resolution by applying dedicated software (5,6), at a certain point this becomes unreliable and unsuccessful.

We hypothesized that the reverse is also true: if the array complexity is reduced by decreasing the number of BAC...
elements (spots) on the chip replacing them with elements of higher complexity (pools of five BAC per spot), more amplification steps are tolerated and the variation in the mean ratio over the whole genome decreases. Reduction of complexity can be accomplished in practice by selecting and pooling a defined number of BACs from the 3.5 k sets that have been produced for instance by the UCSF Cancer Center (1) or the Wellcome Trust Sanger Institute (7). The pooled BACs that together cover a defined region in the genome are then spotted. The selection strategy is based on the following principles: (i) only BACs are selected that so far behaved flawless on the basis of own experience or characteristics listed in the existing databases; (ii) the number of BACs selected for a particular region may be tuned: more BACs means better coverage of the region involved (chromosome arm) and consequently smaller chance to be influenced by a random variation of the mean signal ratio caused by amplification bias; using pools of smaller numbers of BACs on the glass means decreased array complexity with increased probe complexity and greater robustness; (iii) the total number of spots on the array is kept low in order to minimize the hybridization area, which enables the use of a relatively high amount of DNA per unit area.

We have used this strategy to produce mini-arrays of maximally 48 different spots, spotted in 5-fold, of which each spot covers (a pool of evenly distributed flawless five BAC clones) a human chromosome arm; 43 different chromosome arms and control spots (Drosophila BAC, C0t-1 DNA). The performance of these mini-arrays was evaluated on cell lines with known gains and losses, with special emphasis on reducing assay time and amount of target DNA.

We have shown that gains and losses of relatively large genomic regions such as full chromosomes and chromosomal arms are reliably detected in <4 h. Our results indicate that such achievement can be obtained with much lower amounts of target DNA.

**MATERIALS AND METHODS**

**Clone selection for production of mini-arrays**

For clone selection, array-CGH results of thirty normal genomic DNA samples using in house produced 3.5 k BAC arrays (probes kindly provided by the Sanger Centre, Cambridge, UK) were taken (7,8). The normalized mean ratio of triplicate spots of each BAC clone was used to calculate the normalized mean ratio over thirty hybridizations and its corresponding standard deviation. These values were used to select five clones for each chromosome arm, with a mean value close to 1 and a relatively low standard deviation (typically below 0.03). From this set five clones evenly distributed over a given chromosome arm were selected and pooled, see Supplementary Data (clone names and positions are according to freeze November 2004 at Ensembl site: http://nov2004.archive.ensembl.org/). Care was taken not to select clones too close to centromeric or telomeric regions since these are known to contain higher amounts of repetitive sequences and might be involved in segmental duplications.

**Generation of mini-arrays**

BAC DNAs were isolated from bacterial clones, using the Wizard SV 96 Plasmid DNA Purification System (Promega, Leiden, The Netherlands) in combination with the Biomek 2000 Laboratory Automation Workstation (Leiden Genomic Technology Center facilities (LGTC), Leiden, The Netherlands). This DNA purification kit is designed to isolate DNA from plasmids and results in small amounts of DNA (~100 ng DNA from 1 ml culture) when used for BAC isolation. However, in our experience, this system was easier to implement using robotics than usual protocols for BAC DNA isolation. The resulting DNA had low levels of contamination from the host Escherichia coli, and was suitable for DNA amplification and subsequent array production. Amplification of the DNA, spotting on the slides and hybridization procedures were based on protocols optimized by the group of Dr N.P. Carter (Wellcome Trust Sanger Institute, UK). This set of BACs and protocols are described in detail [see ref (7,8)]. For each chromosome arm 20 μl from the secondary DOP PCR amplified product of each selected BAC clone was pipetted together. These clone mixtures were precipitated with sodium acetate and ethanol and dissolved again in spotting buffer at 1 μg/μl concentration in 20 μl volume. Spots were done using an Omnigrid 100 Microarrayer (Genomic Solutions, Ann Arbor, MI) at the Leiden Genome Technology Center facilities (LGTC, Leiden, The Netherlands). A total of 43 clone mixes was spotted, for each of chromosome one arm; as control probes non-human DNA derived from Drosophila melanogaster and amino labelled C0t-1 DNA were spotted.

In addition, one spotting well containing material of four BAC clones derived from the long arm of chromosome 4 and one BAC clone derived from the long arm of chromosome X was prepared; the mixture was spotted and used as a control to show the sensitivity and the dynamic range of detection.

Each clone mixture was spotted five times on the array, (average spot size: 150 μm; spot-to-spot distance: 300 μm) resulting in a mini-array of 4.5 by 5 mm containing 240 spots (0.225cm²).

**Test material**

Performance of the mini-arrays for detecting abnormalities was tested on a set of 8 cell lines with confirmed abnormalities, diagnostically important in clinical cytogenetics. The abnormalities were 47,XX,+13; 47,XX,+18; 47,XX,+21; 46,XX,i(18)(q10); 45,X0; 47,XXX; 47,XY; 49,XXXXY. In addition two solid tumour cell lines were investigated, a Ewing Sarcoma (TC32) and a T cell lymphoma. For all samples also 3.5 k array-CGH analysis was performed for comparison.

**DNA isolation, fluorescent labelling and hybridization**

Test DNA of the cell lines was isolated using a High Pure PCR Template Preparation Kit, (Roche diagnostics, Mannheim, Germany).

Both test and reference DNA were labelled using the BioPrime® DNA Labelling System (Invitrogen, Breda, The Netherlands) with an adapted dNTP mix (7). A total of 100 ng of each DNA were labelled overnight, either with Cy3 or Cy5 dCTP (Amersham, Roosendaal, The Netherlands) with an additional 20 mM of sodium chloride in a total reaction volume of 30 μl. The labelled test and reference DNA were mixed together and precipitated with 37.5 μg C0t-1 DNA.
(Invitrogen, Breda, The Netherlands) and dissolved in 28 μl of hybridization buffer [50% formamide, 2× SSC, 10% dextran sulphate, 0.1% Tween-20, 10 mM Tris–HCl (pH 7.4) and 25 mM EDTA (pH 8.0)] and 2 μl of yeast tRNA (100 μg/μl, Invitrogen, Breda, The Netherlands). Slides were hybridized using 9 by 9 mm frame-seal chambers (BioRad, Veenendaal, The Netherlands) without cover slip. Pre-hybridization was not found necessary and omitted. Hybridization was performed in an airtight sealed humidified chamber, originally a microscope slide storage box with a moist paper on the bottom, containing 50% formamide, 2× SSC, pH 7 and sealed with formamide resistant tape. The hybridization was done overnight in a 37°C incubator on a slowly rocking table.

After hybridization, the frames were removed and slides were washed in a solution of 50% formamide, 2× SSC, pH 7, at 48°C for 15 min, followed by a wash in 2× SSC and 0.1% SDS, at 48°C for 30 min. The slides were transferred to a 0.1 M phosphate buffer, pH 8 with 0.1% Igepal (Sigma-Aldrich Chemie B.V., Zwijndrecht, The Netherlands) at room temperature and washed for 10 min. Then the slides were dipped briefly in MilliQ water and were spun dry.

For dilution experiments 3-fold dilutions of the labelled target DNA were prepared and used with different hybridization times (4, 8, 16 and 24 h). In these experiments, all other conditions were kept the same as described above.

**Data analysis**

The arrays (3.5 k and mini-arrays) were scanned using a GenePix 4100A scanner (Axon Instruments, Union City, CA). Images were processed using GenePix Pro 4.1 software. Pixel intensities for each feature were integrated and median values were determined, and the local background was calculated. For each spot the intensities were corrected (pixel median values) by subtracting the local background (pixel median values) for both wavelengths. The median of the ratios of all spots was calculated and used to normalize all data points. From each feature, the log2 value of the average of the normalized ratios of the five spots was calculated and used to display the data. For each feature (5 spots) the mean value and its standard deviation were calculated. Target values with a standard deviation greater than ±0.2 or signal to background ratios <2 were excluded. A conventional threshold of ±0.25 (on a log2 scale) was used to identify gains and losses.

**RESULTS**

Pools of normal male and female genomic DNA (Promega, Leiden, The Netherlands) were used to do self-self and gender-mismatched hybridizations, followed by gender-mismatched hybridization of genomic DNA of cell lines with known trisomies and monosomies. Spot-to-spot variation within the five replicates was generally below six percent and all expected trisomies and monosomies were picked up with no false positives or false negatives. For potential diagnostic purposes one may consider to filter the data using the minimum/maximum outlier reduction removing the highest and lowest ratio value

![Figure 1](https://academic.oup.com/nar/article-abstract/33/18/e159/2401418)

**Figure 1.** Typical examples of CGH results obtained with mini-arrays (for assay conditions see elsewhere). The log2 value of the ratio is plotted for the p and q arm of each chromosome, starting with chromosome 1p and ending with chromosome Yq. For Figures A–C independent duplo experiments are plotted in the same graph, one experiment is presented with open circles and the other experiment with closed circles. (A) 47,XX,+13 versus 46,XY; (B) 46,XX,i(18)(q10) versus 46,XY; (C) 47,XX,+21 versus 46,XY; (D) average ratio and standard deviation for all chromosome arms was calculated and plotted for 12 independent assays. The total average standard deviation was 0.034.
from the five replicates (9). After the minimum/maximum outlier reduction the spot-to-spot variation between the remaining triplicates dropped in general below 4%. In Figure 1, three representative examples of ratio plots of trisomy DNA versus gender-mismatched normal reference hybridizations are given. To show inter-experimental variation and reproducibility, the ratio plots of the repeated experiment is also given in the same figure (open and closed circles).

For twelve separate hybridizations, the normalized average ratio and its standard deviation was calculated. To generate this data set, normal samples and samples with gains and losses were taken. Therefore to calculate the standard deviation of each chromosome arm representing clone pools, individual data points were omitted from samples with gains and/or losses (trisomy/monosomy). For example in case of chromosome 13 trisomy the ratio value of the chromosome 13q feature was omitted and all other values were used to generate the mean ratio and standard deviation of all other, non-altered clones. The average ratio and variation of these data points per clone is displayed in Figure 1D. The total average standard deviation was 0.034. Signal to background ratios in all standard hybridizations were comparable. As an additional quality measure for hybridization specificity we calculated the ratio of the human BAC signals compared to two different Drosophila BAC clones and to amino-modified COT-1 DNA. In a standard hybridization ratio of the human BAC signal over the Drosophila clones was on average 12.1 and 14.3 for Cy5 and Cy3, respectively. The ratio of the human BAC signals over the spotted human COT-1 was on average 10.8 for Cy5 and 11.4 for Cy3, while the signal versus background ratios were 14.3 and 18.6 for Cy5 and Cy3, respectively. The intensity ratio of the non-human sequences is a measure of hybridization stringency; the intensity ratio of the COT-1 is a measure of the efficacy of repeat blocking and hybridization stringency.

The number of selected BAC clones to generate a pool of any given region of interest will determine the possible influence/sensitivity of an individual clone to detect a rearrangement smaller than a chromosome arm or in the contrary to be influenced by the variation in mean signal ratio due to under- or over-representation of genomic regions after amplifications. For example by using only two clones as a pool, a single copy gain or loss of the genomic region covered by only one of the two clones would result in 50% increase or decrease the ratio of the pooled spot, while using pools of five clones this influence would be only 20%. The number of clones to be selected might vary depending on the goal of the potential application from 5 to 10. The results obtained with the mini-arrays under the conditions described were very similar to the ones obtained by the larger 3.5 k array (Figure 2). All full chromosome arm copy number changes were perfectly identified, while a small interstitial deletion (<4 Mb) on 2q and a complex interstitial deletion pattern on chromosome 13 involving several regions were not reaching the log2 ±0.25 threshold value. By using the 3× SD to set a threshold (−0.109) for the 13q representing clone (−0.098) (Figure 2A and C) the deletion was not detected.

To further test the performance of the mini-array, one BAC clone in a clone mix for the long arm of chromosome 4 was replaced with a BAC clone, unique for chromosome X. One would expect the ratio of this clone mix in a normal-to-normal gender-mismatched hybridization to increase with one fifth of the ratio difference between the X chromosome clones and the autosomal clones. Here using a threshold value based on 3× SD (0.09), the amplification (0.142) was readily detected. Subsequently, for a trisomy X sample assayed against a male reference sample, the ratio of the clone mix with one chromosome X clone is expected to increase accordingly (0.257). Figure 3A and B shows that the results obtained for both experiments well match the theoretical expectation. This proves that one BAC contributes to a fifth of the ratio as it is expected. The use of this information provides the possibility to detect genomic imbalances involving only a part of any

![Figure 2](https://www.nature.com/nar/journal/v33/n18/fig_tab/fig02f.html)
given chromosome arms using threshold values based on the 3× SD of any given clone.

Finally, the amount of labelled target DNA was diluted and the hybridization time was reduced. Although larger variations in normal region became apparent, the abnormalities were still detectable when DNA equivalents of 30 ng of target DNA was used at a hybridization time of only 4 h (Figure 4A), using only 10 ng of target DNA detection identification of alterations was not possible (Figure 4B). However, 10 ng of target DNA was found adequate when the hybridization period was extended to 18 h.

**DISCUSSION AND CONCLUSIONS**

In the past decade genomic arrays have been successfully used to detect genomic amplifications and deletions in virtually all types of human tumours. For a few, the abnormalities found are characteristic and allow identification of the tumour type, or they are used prognostically or provide extra rational for selection of proper therapy. The majority of abnormalities found however, is strongly case dependent and occurs all over the genome, albeit that some regions are more frequently involved than others. Array-CGH has been an extremely valuable discovery tool, which in many cases has elucidated amplifications and deletions that represent the critical steps in tumorigenesis of many tumours. A significant improvement in detecting smaller amplifications and deletions was made with the introduction of spotted high-density BAC arrays (35 k), or high-density oligonucleotide arrays to detect single nucleotide polymorphisms (SNPs). The higher resolution data that is produced with these devices is ongoing, and the number of clinically relevant gene amplifications and deletions is increasing.

When the aim is diagnosis rather than discovery, the situation changes. Although genomic arrays obviously can be used diagnostically, there are many suitable alternatives, the choice of which depends on a number of variables. First, the multiplicity, that is the number of gene regions that have to be investigated. When small, interphase cytogenetics is a good alternative, for instance the use of a two colour fluorescence in situ hybridization (FISH) test to demonstrate Her2/Neu amplification for selection of Herceptin®/C210 treatment for breast cancer. When the number of targets is higher, for instance up to 100, other technologies are of interest. Examples are bead-based platforms or multiplex ligation-dependent probe amplification (MLPA) (10). As this test format requires a 5th to 10th of labelled DNA to hybridize, consequently the use of a mini-array system could reduce the labelling related costs to the same extent. In addition, multiple mini-arrays could be spotted on one array-slide (three in our system, but four or more could be adapted) resulting in a proportional reduction of slide related costs per assay. These considerations would
favour the use of mini-array in diagnostic procedures or even in a kind of pre-testing of samples (information about sample quality) before one would use an array with higher resolution and higher related costs.

Here we show a strategy to match the performance of (home) spotted BACs to the goal for which they are used. Earlier, attempts have been made to pool chromosomal painting probes, and produce mini-array to detect aneuploidies, but with limited success (11). The complexity of painting probes is much higher (the complexity of a spot element is a full chromosome) than the BAC clones (one spot element containing five clones distributed over a given chromosome arm) used here. Also, that approach lacked the possibility of selecting the BAC clones with the best performance, and in the desired number.

The mini-arrays show excellent performance in detecting aneuploidy of a given chromosome, a feature that for instance flow cytometry does not offer. More importantly, informative results can be obtained using low amounts of target DNA, and easily within one day, DNA isolation and labelling included. Low amounts of target DNA are important when only limited material of small tumours or their metastases is available. When one accepts a limited resolution of one chromosomal arm, mini-arrays may also be applied for genomic profiling of disseminated tumour cells that are isolated from bone marrow by micromanipulation of by laser microdissection; a situation where the amount of DNA is very restricted. For all data shown in Figures 1–4 we used exclusively random primed labelled DNA and did not apply techniques such as linker-adaptor PCR or φ-29 phage amplification that are successfully used to amplify DNA isolated from single cells (4). In combination with such amplification techniques, a highly interesting application of this mini-array strategy could be detection of chromosomal aneuploidies in the context of preimplantation genetic diagnosis (PGD). PGD of single blastomeres to detect aneuploidies has been shown to increase the success rate in genetic diagnosis (PGD). PGD of single blastomeres to detect chromosomal aneusomies in the context of preimplantation genetic diagnosis (PGD) offered a highly interesting application of this mini-array strategy, which could be used to detect chromosomal aneuploidies. PGD of single blastomeres to detect chromosomal aneusomies has been shown to increase the success rate in genetic diagnosis (PGD). PGD of single blastomeres to detect aneuploidies has been shown to increase the success rate in genetic diagnosis (PGD). PGD of single blastomeres to detect aneuploidies has been shown to increase the success rate in genetic diagnosis (PGD). PGD of single blastomeres to detect aneuploidies has been shown to increase the success rate in genetic diagnosis (PGD).

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

Funded in part by ‘Doelmatigheid’ grant, Leiden University Medical Center (2002/2003), Leiden, The Netherlands.

Conflict of interest statement. None declared.

REFERENCES

1. Pinkel,D., Segraves,R., Sudar,D., Clark,S., Poole,I., Kowbel,D., Collins,C., Kuo,W.L., Chen,E., Zhai,Y. et al. (1998) High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays. Nature Genet., 20, 207–211.
2. Ishkanian,A.S., Molloff,C.A., Watson,S.K., DeLeeuw,R.J., Chi,B., Coe,B.P., Snoijers,A., Albertson,D.G., Pinkel,D., Marra,M.A. et al. (2004) A tiling resolution DNA microarray with complete coverage of the human genome. Nature Genet., 36, 299–303.
3. Kraus,J., Pantel,K., Pinkel,D., Albertson,D.G. and Speicher,M.R. (2003) High-resolution genomic profiling of occult micrometastatic tumor cells. Genes Chromosomes. Cancer, 36, 159–166.
4. Langer,S., Geigl,J.B., Gangnus,R. and Speicher,M.R. (2005) Sequential application of interphase-FISH and CGH to single cells. Lab Invest, 85, 582–592.
5. Eilers,P.H. and de Menezes,R.X. (2005) Quantile smoothing of array CGH data. Bioinformatics, 21, 1146–1153.
6. Jong,K., Marchiori,E., Meijer,G., Vaart,A.V. and Ylstra,B. (2004) Breakpoint identification and smoothing of array comparative genomic hybridization data. Bioinformatics, 20, 3636–3637.
7. Fiegler,H., Carr,P., Douglas,E.J., Burford,D.C., Hunt,S., Scott,C.E., Smith,J., Verdie,D., Gorman,P., Tomlinson,I.P. and Carter,N.P. (2003) DNA microarrays for comparative genomic hybridization based on DOP-PCR amplification of BAC and PAC clones. Genes Chromosomes. Cancer, 36, 361–374.
8. Knijnenburg,J., Szuhai,K., Gillay,J., Molenar,L., Sloos,W., Poot,M., Tanke,H.J. and Rosenberg,C. (2005) Insights from genomic microarrays into structural chromosome rearrangements. Am. J. Med. Genet. A, 132, 36–40.
9. Wessendorf,S., Fritz,B., Wrobel,G., Nessling,M., Lampel,S., Goettel,D., Kuepper,M., Joos,S., Hopman,T. and Kokocinski,F. et al. (2002) Automated screening for genomic imbalances using matrix-based comparative genomic hybridization. Lab Invest., 82, 47–60.
10. White,S.J., Breuning,M. and den Dunnen,J.T. (2004) Detecting copy number changes in genomic DNA: MAPH and MLPA. Methods Cell Biol., 75, 751–768.
11. Hu,D.G., Webb,G. and Hussey,N. (2004) Aneuploidy detection in single cells using DNA array-based comparative genomic hybridization. Mol. Hum. Reprod., 10, 283–289.
12. Munne,S. (2003) Preimplantation genetic diagnosis and human implantation–a review. Placenta, 24, S70–S76.
13. Baert,E.B., Martini,E. and Van Opstal,D. (2004) Screening for aneuploidies of ten different chromosomes in two rounds of FISH: a short and reliable protocol. Prenat. Diagn., 24, 955–961.
14. Wilton,L., Voullaire,L., Sargeant,P., Williamson,R. and McBain,J. (2003) Preimplantation aneuploidy screening using comparative genomic hybridization or fluorescence in situ hybridization of embryos from patients with recurrent implantation failure. Fertil. Steril., 80, 860–868.
15. Verlinsky,Y. and Kuliev,A. (2003) Preimplantation diagnosis for aneuploidies using fluorescence in situ hybridization or comparative genomic hybridization. Fertil. Steril., 80, 869–870.