Microarray CGH

Ben Beheshti, Paul C. Park, Ilan Braude, Jeremy A. Squire

Ontario Cancer Institute, Princess Margaret Hospital, University Health Network, and Departments of Laboratory Medicine and Pathobiology, and Medical Biophysics, Faculty of Medicine, University of Toronto, Ontario, Canada.

From: B. Beheshti, P.C. Park, I. Braude, J.A. Squire. "Microarray CGH". *In:* Y.-S. Fan (Ed.), Molecular Cytogenetics: Protocols and Applications: Humana Press, 2002.

*Corresponding Author:

JA Squire, Ph.D.
Ontario Cancer Institute
Division of Cellular and Molecular Biology
610 University Ave. Room 9-721
Toronto, Ontario, Canada
M5G 2M9
E-mail: jeremy.squire@utoronto.ca
Fax 416-920-5413

Table of Contents

1. Introduction

1.1 cDNA Array CGH

1.1.1. Application of cDNA Array CGH to Cancer Genomics

1.1.2. Current Limitations of cDNA Array CGH

1.2 Array CGH

1.2.1. Application of Array CGH to Cancer Genomics

1.2.2. Applications in Other Fields

1.2.3. Current Issues with Genomic DNA-based Array Fabrication

1.2.4. Current Accessibility to Genomic DNA-based Arrays

1.2.5. Commercial Sources of Genomic DNA-based Arrays

1.3. Detection and Analysis

1.3.1. Image Acquisition

1.3.2. Fluorescence Quantification and Ratio Analysis

1.3.3. The Role of Bioinformatics in Microarray CGH

2. Materials

2.1. cDNA Array CGH

2.1.1. Array Preparation

2.1.2. Probe Preparation by Random Primer Labeling of Genomic DNA

2.1.3. Probe Denaturation and Hybridization

2.1.4. Washes

2.2. Array CGH

2.2.1. Array Preparation

2.2.2. Probe Preparation by Nick Translation of Genomic DNA

2.2.3. Probe Denaturation and Hybridization

2.2.4. Washes

3. Methods

3.1. cDNA Array CGH

3.1.1. Array Preparation

3.1.2. Random Primer Labeling of Genomic DNA

3.1.3. Probe Denaturation and Hybridization

3.1.4. Washes

3.2. Array CGH

3.2.1. Array Preparation

3.2.2. Probe Preparation by Nick Translation of Genomic DNA

3.2.3. Probe Denaturation and Hybridization

3.2.4. Washes

4. Notes

5. References

6. Figure Legends

6.1. Figure 1

6.2. Figure 2

6.3. Figure 3

6.4. Figure 4

6.5. Figure 5

1. Introduction

Comparative genomic hybridization (CGH) to metaphase chromosome targets (1,2) has significantly contributed to our understanding of the cancer cytogenetics of more complex malignancies such as the solid tumours (*see* chapter 9; reviewed in (3,4)). This molecular cytogenetics-based technique (hereafter referred to as "chromosome CGH") is capable of defining genome-wide DNA copy number imbalances in sample cells relative to a normal reference in a single experiment. Chromosome CGH has greatly increased our understanding of tumour biology and progression since the minimal recurrent regions of chromosomal gain and loss are likely to contain novel oncogene(s) and tumour suppressor gene(s) respectively.

Limitations of Chromosome CGH

The unique advantage of chromosome CGH is its whole-genome screening capability which is significantly faster and less laborious than low-throughput methods for examining single-target dosage changes such as Southern analysis, PCR, and fluorescence *in situ* hybridization (FISH). Chromosome CGH is now a well-established molecular cytogenetic method, but there are two technical limitations that restrict its usefulness as a comprehensive screening tool. First, because the target DNA within the chromosome is highly condensed and supercoiled, the resolution for determining copy number changes is no less than 10 Mb for loss (1). For copy number gains, the minimal detectable size is probably no less than 2 Mb, which is a function of both amplicon size and copy number (1,5). This resolution, while capable of providing a starting point for positional cloning studies, will still encompass too many genes to precisely localize a sequence of interest. Second, the analysis of the images obtained following chromosome

CGH is only partly automated and experienced cytogeneticists must identify each chromosome to determine regions of imbalances.

Microarray CGH: Application of Microarray Technology to CGH

Recent developments in microarray methods have circumvented some of the limitations of chromosome CGH. Complementary DNA (cDNA) microarray technology, realized through advances in the Human Genome Project (HGP) as well as robotic arraying technology on glass slides, has facilitated high-throughput analysis of differential gene expression in tumours (6-8). An emerging platform that addresses the shortcomings of chromosome CGH couples the technique to microarray expression technology, and is generally referred to as "microarray CGH". Instead of using metaphase chromosomes, CGH is applied to arrayed short sequences of DNA bound to glass slides (herein defined as the "targets" for hybridization) and probed with genomes of interest (herein defined as the "probe") (*see* Note 1). With sufficient representation on the microarray, this system significantly increases resolution for localizing regions of imbalance. Furthermore, just as with expression microarray screening, analysis is straightforward and automated. Two technology platforms have recently been published: 1) cDNA-based array CGH (9,10); and 2) genomic DNA-based array CGH (also referred to as "matrix CGH" and "array CGH") (11,12). This chapter will provide an overview of the currently published methods, but readers should be aware that microarray CGH is an emerging technology and there are likely to be continual refinements to the protocols described below (see Note 2).

1.1. cDNA Array CGH

Microarray CGH using cDNA targets (hereafter referred to as "cDNA array CGH") was first described by Pollack et al. (9). This platform makes use of conventional cDNA microarrays, normally employed in expression screening, for examining genomic copy number imbalances. As depicted in Figure 1, this has the advantage that duplicate arrays may be used in parallel to provide a comprehensive overview of both expression and gene copy number change in a tissue (9). The increasing availability of a variety of different cDNA microarray expression formats means that modification of protocols to interrogate these cDNA targets by CGH is immediately accessible for high-throughput analysis of gene dosage changes.

1.1.1. Application of cDNA Array CGH to Cancer Genomics

Pollack et al. examined breast cancer cell lines and tissues using a 3,360 feature microarray by cDNA array CGH (9). With optimization, they demonstrated that the technique was capable of detecting copy number gains and single deletion losses. Analysis of the tumours and cell lines showed that not all amplified genes were overexpressed, nor were most highly overexpressed genes amplified; however, a subset of the genes, including *ERBB2*, were observed to be both amplified and overexpressed. They proposed that these genes might be important mediators of the tumour initiation and progression.

The utility of cDNA array CGH for detecting gene amplifications was recently shown by Heiskanen et al. (13). In this study, cell lines with known gene amplifications were used to establish the sensitivity limits of the technique. In contrast to the protocol used by others (9,10), genomic DNA is biotin-labeled and a tyramide amplification

protocol (14) is employed (13). Progressive dilution from 100% to 2% of genomic DNA from the neuroblastoma cell line NGP with normal DNA during labeling corresponded to decreasing *MYCN* signal intensity on the microarray. Amplifications of 5-fold and greater were readily detected by this method, and at 2% dilution *MYCN* intensity was observed at 2.5-fold relative to other non-amplified genes. However, the main limitation of this method is its inability to allow two-colour CGH and thus necessitates the use of two microarrays (test, control) per experiment.

Recently, we have demonstrated the suitability of cDNA array CGH for gene amplification screening of patient samples (10). In this study, the *MYCN* (chromosome region 2p24) amplification status in neuroblastoma patients and cell lines was confirmed by cDNA array CGH on a high-density 19,200 feature microarray. In the cell line IMR32, cDNA array CGH confirmed a recently described co-amplified oncogene, *MEIS1* (15,16). Importantly, the technique was able to distinguish three tumour genotypes in patient samples not previously described (Figure 2). This study demonstrates not only the high-throughput advantage of examining thousands of genes by cDNA array CGH over conventional methods such as FISH and Southern analyses, but also the increase in resolution in contrast to chromosome CGH.

In another study by Pei et al. (17), the increased resolving power of cDNA array CGH for delineating amplicon boundaries was demonstrated in pediatric carcinomas. This work clearly shows the limited resolution of chromosome CGH when contrasted to cDNA array CGH. These results are depicted in Figure 3.

1.1.2. Current Limitations of cDNA Array CGH

There are at least three limitations to current cDNA array CGH methods. Firstly, target cDNA sequences are of low complexity in content in comparison to genomic sequences, lacking intronic and other non-transcribed elements such as repetitive DNA and control sequences. Thus many regions of the genome being interrogated will not hybridize with uniform efficiency so that the specificity of the technique may be low or poorly reproducible. Secondly, target cDNA sequences are typically only 0.5-2 kilobases in size (9,10,13). This is on a scale of many orders of magnitude smaller than the smallest chromosome, and 1-2 orders of magnitude smaller than genomic insert sequences in bacterial artificial chromosomes (BACs), P1-derived artificial chromosomes (PACs), and cosmids. Although this may be suitable for expression mapping by microarray where the probe is comparable in size, reduced signal sensitivity may become a concern when using labeled genomic probes. Although Pollack et al. (9) describe detection of both copy number gains and losses by cDNA array CGH, it is likely that genomic DNA-based arrays are more robust for detection of single copy changes, including copy losses. Finally, a last issue with cDNA microarray technology, and therefore also with cDNA array CGH, is that currently there is a significant number of gene misannotations in the commercially available clone sets (18). This may take the form of wrongly identified sequences, incorrect chromosomal locations or even the complete absence of human sequences in the cDNA targets (eg. due to clone contamination, heterologous sequences). In practical terms this manifests as inconsistent results or findings that cannot be substantiated when other methods are applied. To eliminate this shortcoming, commercial sources of clone sets and many institutions with

array fabrication capabilities are sequence-confirming their clone sets. Overall, these limitations contribute to the high rate of false positive (15%) and false negative (15%) results reported for this technique (9).

1.2. Array CGH

The second microarray CGH platform (hereafter referred to as "array CGH") uses genomic DNA sequences as targets on the microarray. Array CGH was first established by Solinas-Toldo et al. (11), and further refined by Pinkel et al. (12). As described in these studies, the DNA targets for the microarray can be derived from genomic clones including yeast artificial chromosome (YAC; 0.2-2 Mb in size), BAC (up to 300 kb), P1 (~ 70-100 kb), PAC (~ 130-150 kb), and cosmid (~ 30-45 kb), and are of several orders of magnitude smaller than chromosome targets. This decrease in target size increases the resolution of copy number imbalance detection over chromosome CGH (Figure 4). Given the differences in the structural complexity in the target DNA with respect to chromosome CGH, modifications to the hybridization conditions are necessary (11,12). The advantage of array CGH over cDNA array CGH is that there is more uniformity in hybridization and subsequent signal fidelity because the DNA targets have a greater complexity and coverage, containing intronic and other non-transcribed genomic sequences.

1.2.1. Application of Array CGH to Cancer Genomics

To date, several groups have published results using array CGH (11,12,19-25). Pinkel et al. (12) detected genomic imbalances within a sub-band of chromosome 20 in breast cancer that had failed to be observed using chromosome CGH. Using array CGH,

precise genomic mapping of the position of amplicon boundaries within 20q13.2 was performed (19). This allowed *CYP24* to be localized within the minimal amplified region, identifying it as a new candidate oncogene in breast cancer (19). In another study, array CGH was used to examine neurofibromatosis type 2 (NF2) patients and determined the extent and frequency of deletions around the *NF2* locus on chromosome 22q (20). This microarray was constructed from a 7 Mb tiling path of 104 BAC and PAC genomic clones around *NF2*, and included smaller cosmids for mapping copy number changes at higher resolution. Both single copy losses and homozygous deletions were detectable in the patient samples by this system (Figure 5). Further refinements have permitted retrospective analysis using genomic DNA from archival samples. Daigo et al. (21) have adapted array CGH for amplicon profiling of laser capture microdissected (Arcturus, Mountain View, CA; http://www.arctur.com) formalin-fixed paraffin-embedded tumour samples, using degenerate oligonucleotide-primed (DOP)-PCR (26) for whole genome amplification of the extracted DNA.

1.2.2. Applications in Other Fields

Microarray CGH is a versatile technique that may be used to examine genetic disorders other than cancer. A recent study by Geschwind et al. (23) demonstrated the use of array CGH for investigating the molecular basis of laterality of the human cerebral hemispheres. Gene dosage changes in patients with Klinefelter's syndrome (karyotype: XXY) were examined with a DNA microarray constructed with cosmids covering the pseudoautosomal region of the sex chromosomes, and findings were correlated with anomalous dominance and other cognitive or behavioural phenotypes.

1.2.3. Current Issues with Genomic DNA-based Array Fabrication

Although array CGH still has some limitations, most of these relate to array production and will be addressed as the technology matures. While modifications to existing array fabrication systems are possible, current production limitations are mainly associated with difficulties in automating batch preparation DNA from genomic clones. For example, published array CGH studies involve the use of laborious DNA extraction methods such as maxi prep kits (Qiagen) and phenol/chloroform extractions from genomic clones (11,12). However, commercially available batch extraction kits (eg. R.E.A.L. System™, Qiagen) from genomic clones coupled with DOP-PCR may aid in automation (*see* Note 3). A second difficulty is related to the generation of adequate amounts of DNA for batch microarray production. While cDNA expression clones have universal primer sites amenable to large scale PCR synthesis of genes and expressed sequence tags for subsequent purification and arraying, the same is not true for genomic clones. In addition, the larger genomic inserts require long PCR which has more exacting amplification conditions (27) and these may be confounded by the presence of repetitive DNA elements in template sequences. The third difficulty is the viscosity of large size genomic sequences in solution that may cause clogging of spotting pins of some arrayers, although new split pin designs may circumvent this problem (28). Finally, as with the cDNA clone sets, there is also the concern that a small but significant number of commercially available genomic clones are misannotated in their localization (eg. due to source clone plate contamination, mislabeling). Currently the solution is to FISH-confirm cytogenetic mappings of clones used for array CGH, although this is not a trivial task when dealing with tens or hundreds of genomic clones. The BAC/PAC resources

(http://www.chori.org/bacpac/), further described in chapter 27, is an ongoing project to FISH-map all clones (25) that will largely alleviate this problem.

1.2.4. Current Accessibility to Genomic DNA-based Arrays

While cDNA microarrays can be obtained both commercially and from array fabrication core facilities within research institutions, array CGH is not yet immediately accessible to most researchers. At present, scientists wanting to study a chromosomal region of interest by array CGH will require custom array production. Progress in the HGP has facilitated construction of a tiling path of genomic clones that cover chromosomal loci of interest (eg. MapViewer resource at the National Center for Biotechnology Information: http://www.ncbi.nlm.nih.gov/). While the associated costs of genomic DNA-based microarray production are not practical for individual research laboratories, it is likely that institutional core microarray facilities will be able to modify production to address this need. Conceivably, the post-HGP era will facilitate production of whole genome arrays (29), and even higher-resolution chromosome-specific and chromosome band-specific microarrays. Notably, the first high-density whole genome microarray (approximately 2,000 BAC clones) was recently introduced (25) and demonstrated its ability to precisely delineate genome-wide segmental aneuploidy breakpoints in tumour cells.

1.2.5. Commercial Sources of Genomic DNA-based Arrays

An alternative to custom arraying of genomic targets may be to obtain commercially available microarrays. One such system is produced by Vysis Corporation (http://www.vysis.com), called the GenoSensor System™. The AmpliOnc I™ array from Vysis contains BAC, PAC, and P1 genomic clones from 59 known oncogenes

spotted in triplicate (30), and has been used by groups studying breast cancer (21) and glioblastoma multiforme (24). This microarray complements their GenoSensor™ microarray reader and analysis software package. The next generation genomic microarray from Vysis will comprise 250-300 features, including genomic clones from the AmpliOnc I™ array, subtelomeric regions of all chromosomes, major tumour suppressor genes, and major microdeletion syndrome loci (30). Recently, Spectral Genomics (http://www.spectralgenomics.com/) has produced a commercially available whole-genome human BAC microarray kit. The current generation microarray is spotted in duplicate with 1003 human BAC clones, spaced at regular intervals along the genome, giving an effective resolution of 3 Mb for defining genomic aberrations. It is expected that both higher resolution (1 Mb and less) human and mouse BAC microarrays become available for purchase in the near future.

1.3. Detection and Analysis

Analysis of microarray CGH involves three components, namely: 1) image acquisition; 2) quantification of fluorescence intensity; and 3) interpretation. These can be accomplished using the system developed for expression microarrays with minimal or no modification.

1.3.1. Image Acquisition

Image acquisition for microarray CGH requires systematic scanning of all gridded features on the microarray. Commercially available microarray scanners are typically laser-based scanning systems that can acquire the two differential wavelengths sequentially (eg. Packard BioScience, http://www.packardbiochip.com) or

simultaneously (eg. Virtek Vision Inc., http://www.virtek.ca; Axon Instruments Inc., http://www.axon.com). Alternatively, resources for the development of in-house microarray scanning systems are also available (eg. http://brownlab.stanford.edu/; (31)). The technical details underlying these systems are specific to the hardware package, and are beyond the scope of this chapter.

1.3.2. Fluorescence Quantification and Ratio Analysis

Software for fluorescence quantification and ratio analysis of gridded spots is usually included with the scanner hardware. Alternatively, there are less sophisticated softwares publicly available (eg. ScanAlyze: http://rana.stanford.edu/; (32)). Quantified fluorescence intensities requires normalization and establishment of the fluorescence ratio baseline. Often, microarray features are spotted in duplicate or triplicate for assessing result reproducibility. For array CGH, inclusion of genomic clones onto the microarray from regions that are known not to be involved in copy number change are recommended as internal controls for these purposes. In addition, parallel experiments in which differentially labeled normal genomic DNA is compared against itself can serve to establish the specificity of the system. Overall, there is an obvious need for statistical analysis of the conformity of the results (33). Global normalization approaches such as those used in expression microarray experiments may also be used for establishing baseline thresholds (10,34).

Previous reports indicate that the relationship between the fluorescence ratio and copy number changes (1,9,11,12) deviates from linearity at low copy numbers. For this reason, it is important for users to independently establish this relationship for

interpretation of CGH results and to confirm imbalances by direct FISH analysis of tissue sections.

1.3.3. The Role of Bioinformatics in Microarray CGH

As representation on the microarrays increases in density, data storage (35) and bioinformatics will become an important aspect of the CGH analysis. In addition, the increase in resolution will make the task of identifying consensus regions of genomic imbalance amongst samples more challenging. Overall, this will necessitate datamining techniques that can handle many data points on multiple dimensions between experiments. Moreover, for cDNA array CGH, *in silico* determination of chromosomal localisations of cDNA targets is essential for providing a comprehensive ideogram-type schematic of chromosomal copy number changes (Figure 3) (10). As microarray CGH technology becomes more prevalent, more standardized informatics and analysis tools will appear.

Acknowledgments

The authours are grateful to Paula Marrano, Bisera Vukovic, Monique Albert, and Pascale Macgregor for critical reading of the manuscript. This work was supported by the National Cancer Institute of Canada, with funds from the Canadian Cancer Society. PCP was supported by the AFUD and the Imclone Systems, Inc.

2. Materials

2.1. cDNA array CGH

2.1.1. Array preparation

1. 20X sodium saline citrate (SSC): Dissolve 175.32 g of NaCl, 88.23 g of sodium citrate-$2H_2O$ in 1 L water, titrate to pH 7.0. Store at room temperature.
2. cDNA microarray. Store in dessicator at room temperature.
3. Blocking solution: 3% BSA, 4X SSC, 0.1% Tween-20. Store at –20ºC.
4. Glass coverslips.

2.1.2. Probe preparation by random primer labeling of genomic DNA

1. High molecular weight genomic DNA.
2. *EcoR*I or *Dpn*II (New England Biolabs).
3. Qiaquick PCR purification kit (Qiagen).
4. BioPrime labeling kit (Gibco BRL). Store at –20ºC.
5. dNTP mixture: 4.8 mM each of dATP, dGTP, dTTP.
6. 2.4 mM dCTP.
7. 1 mM Cy5-dCTP, Cy3-dCTP (Amersham). Store in the dark at –20ºC.
8. Microcon 30 filter (Amicon).
9. Yeast tRNA (Gibco BRL). Store at –80ºC.
10. Poly(dA-dT) (Sigma). Store at –20ºC.
11. Cot-1 DNA (Gibco BRL).

12. Hybridization buffer: 3.4X SSC and 0.3% SDS. Prepare fresh per experiment.

2.1.3. Probe denaturation and hybridization

1. Rubber cement.
2. Hybridization oven.

2.1.4. Washes

1. Heated water bath.
2. Coplin jars.
3. Slide centrifuge.

2.2. Array CGH

2.2.1. Array preparation

1. DNA extracted and purified from genomic clones.
2. Maxiprep DNA extraction kit (Qiagen).
3. Glass slides.
4. Glass capillary tubes or robotic arrayer.
5. Blocking solution: 10 µg/µL salmon sperm DNA (Life Technologies) in 50% formamide (Gibco BRL), 10% dextran sulphate, 2X SSC, 0.2% SDS, 0.2% Tween-20. Store at –20ºC.

2.2.2. Probe preparation by nick translation of genomic DNA

1. High molecular weight genomic DNA.
2. DNA polymerase I (Roche).
3. DNase I (Gibco BRL).

4. 10X Cy3 dNTPs: 0.1 mg/mL BSA (Sigma), 0.1 M β-mercaptoethanol (Sigma), 0.5 M Tris-HCl, 50 mM $MgCl_2$, 0.08 mM Cy3-dCTP (Amersham), 0.2 mM dATP, 0.12 mM dCTP, 0.2 mM dTTP, 0.2 mM dGTP; dissolved in water. Store in the dark at –20ºC.
5. 10X Cy5 dNTPs: 0.1 mg/mL BSA, 0.1 M β-mercaptoethanol, 0.5 M Tris-HCl, 50 mM $MgCl_2$, 0.08 mM Cy5-dCTP, 0.2 mM dATP, 0.12 mM dCTP, 0.2 mM dTTP, 0.2 mM dGTP; dissolved in water. Store in the dark at –20ºC.
6. DNAse I dilution buffer: 50 mM Tris-HCl, 5 mM $MgCl_2$, 1 mM β-mercaptoethanol, 100 µg/mL BSA; dissolved in water. Store at –20ºC.
7. DNA size standard ladder (eg. *Hind*III ladder).
8. 0.3M Ethylenediaminetetracetic acid (EDTA) (Gibco BRL).
9. Sephadex G50 spin column (Amersham).
10. Cot-1 DNA (Gibco BRL).
11. Hybridization buffer: 50% formamide, 10% dextran sulphate, 2X SSC, 2% SDS. Store at –20ºC.

2.2.3. Probe denaturation and hybridization

1. Hybridization oven.

2.2.4. Washes

1. Heated water bath.
2. 0.1 M sodium phosphate buffer.
3. NP-40 (Vysis).

3. Methods

3.1. cDNA array CGH

3.1.1. Array preparation

1. Block cDNA microarray under a glass coverslip for 1 hour at 37ºC with blocking solution prior to hybridization with denatured probe (*see* Note 4).

3.1.2. Random primer labeling of genomic DNA

1. 2 µg each of high molecular weight tumour and normal genomic DNA is separately digested with *Dpn*II for 1-1.5 hours (*see* Notes 5-7). The digestion products are purified (Qiaquick PCR kit), vacuum-dried, and resuspended in 25 µL of water.
2. Random primer labeling is performed using the Bioprime Labeling kit, according to manufacturer's instructions, with modifications. Denature the DNA and 20 µL Random Primers (included in kit) at 100ºC for 5 minutes. Immediately chill on ice, and add 2.5 µL dNTPs, 1.25 µL dCTP, 1 µL Cy5/Cy3-dCTP, and 1 µL Klenow fragment (included in kit). Incubate at 37ºC for 90 minutes.
3. Combine Cy3- and Cy5-labeled products and load onto a microcon 30 filter. After centrifuging at 2,000 g for 10 minutes, check the sample reservoir for the presence of labeled product (purple colour). Add directly to the sample reservoir 30 µg Cot-1 DNA, 100 µg yeast tRNA, and 20 µg poly(dA-dT), and centrifuge for 20 minutes at 5,000 g. To recover the sample, add 15 µL

hybridization buffer, and invert microcon filter into a fresh collection tube and centrifuge for 1 minute at 16,000 g.

3.1.3. Probe denaturation and hybridization

1. Denature the probe at 100ºC for 90 seconds in heated water bath or PCR machine. Chill probe on ice, and allow probe to preanneal at 37ºC for 0.5-1 hour.
2. The probe is added to the microarray, covered with a glass coverslip and sealed with rubber cement. Hybridization is at 65ºC for 16-20 hours in a moist chamber humidified with hybridization buffer (*see* Notes 4 and 8).

3.1.4. Washes

1. The cDNA microarray is washed at 65ºC (*see* Note 8) for 5 minutes in 2X SSC, 0.03% SDS, followed by successive washes in 1X SSC and 0.2X SSC at room temperature (5 minutes each).
2. The microarray is centrifuged at low speed (50 g) for 5 minutes to dry.

3.2. Array CGH

3.2.1. Array preparation

1. Genomic clones (BACs, PACs, cosmids, etc.) are grown with appropriate antibiotic and isolated using commercially available maxi kits. Typical yield is tens of micrograms of DNA. Standard protocols using phenol/chloroform may be used to further purify the DNA (*see* Note 3).
2. Size and quality of DNA is assessed by 1% agarose gel electrophoresis, and quantified with a UV spectrophotometer.

3. This target DNA is sonicated to 1.5-15 kb fragments, precipitated, diluted to appropriate concentrations and spotted down on glass slides in a clean environment with capillary tubes at approximately 200–400 μm diameter spots (*see* Note 9).
4. Arrays are preannealed for 1 hour at 37ºC with 20 μL blocking solution under a glass coverslip in a hybridization chamber (*see* Notes 4 and 10).

3.2.2. Probe preparation by nick translation of genomic DNA

1. 2 μg each of high molecular weight tumour and normal genomic DNA (*see* Note 6) is separately labeled by nick translation. The reaction mixtures are as follows:

   A) Cy3 reaction (to total 100 μL with water):

      i. Tumour genomic DNA: 2 μg
      ii. 10X Cy3 dNTPs: 10 μL
      iii. DNA polymerase I: 1 μL
      iv. DNase I (*see* Note 11)

   B) Cy5 reaction (to total 100 μL with water):

      i. Normal genomic DNA: 2 μg
      ii. 10X Cy5 dNTPs: 10 μL
      iii. DNA polymerase I: 1 μL
      iv. DNase I (*see* Note 11)

2. The labeling reaction proceeds for 1.5 hours at 16ºC (refrigerated water bath or PCR machine), following which the reaction mixtures are put on ice.

3. The size of the labeled product is assessed by 1% agarose gel electrophoresis (*see* Note 12). Optimum fragment length for CGH is 500-2,000 base pairs. If the size range is too large, reaction mixtures are returned to 16ºC with additional DNase I and polymerase I to incubate further.
4. Labeling reaction is stopped with addition of 0.1 volume 0.3M EDTA.
5. Unincorporated nucleotides are removed from the labeling mixtures using a Sephadex G50 spin column.
6. Labeled products are mixed together, supplemented with 50 μg Cot-1 DNA, and precipitated with 0.1 volume 3M sodium acetate and 2 volumes cold 100% ethanol. Precipitate is rinsed with 70% ethanol and air dried, then redissolved in 20 μL hybridization buffer.

3.2.3. Probe denaturation and hybridization

1. Denature probe for 5 minutes at 75ºC, and allow preannealing of the probe for 0.5-1 hour at 37ºC to ensure sufficient blocking of repetitive elements.
2. Apply the probe to the microarray after preannealing of the microarray is completed, cover with glass coverslip and seal with rubber cement. Arrays are hybridized for 24 hours at 37ºC in a chamber humidified with hybridization buffer (*see* Note 4).

3.2.4. Washes

1. Arrays are washed at 55ºC in 50% formamide, 2X SSC pH 7.0 (3X, 10 minutes each), then in 0.1 M sodium phosphate buffer with 0.1% NP-40 pH 8 at room temperature, 5-10 minutes.

2. Drain excess liquid and mount slide in DAPI/Antifade under a glass coverslip.

4. Notes

1. Controversy exists in establishing a standard nomenclature. Although the term "probe" correctly refers to the known nucleic acid sequence tethered on the microarray while "target" is the unknown sequence in the sample (36), for the sake of conformity this chapter is following the convention used by all current microarray CGH publications.
2. For updated protocols to those listed within this chapter, please visit http://www.utoronto.ca/cancyto/.
3. Until automated and practical batch methods are developed, many groups are using maxi kits for obtaining target DNA for genomic DNA-based microarrays. This is a labor- and time-intensive process that needs repeating when the target DNA is exhausted over multiple arrayings. If purified target DNA is available (at least several hundred nanograms template, from either maxi or mini preps), DOP-PCR (26) may be used to ensure an indefinite supply of target DNA.
4. It is very important that the microarray does not dry during any hybridization step. Ensure that the hybridization chamber remains humidified with hybridization buffer to prevent evaporation of the probe or blocking mixture. If the microarray does dry, the results are invariably unusable.
5. The protocol herein is optimized for cDNA microarrays with approximately 3,500 features arrayed over an area of approximately 18 x 16 mm^2 (9,10). The amount of DNA, as well as the final hybridization volume, must be scaled up when using higher density microarrays covering a larger spotting area (10).

6. As expected, the size and purity of the unlabeled genomic DNA is very important for obtaining high quality results using microarray CGH. Low quality DNA used in labeling can result in high background and low signal intensity on the microarray. The protocol stated herein is optimized for genomic DNA extracted from fresh tissues.
7. The choice of restriction enzyme for digestion is important for labeling efficiency. It has been noted that decreasing the average fragment size prior to labeling may increase labeling efficiency (9). This has to be balanced against excessive digestion producing fragments that are too small to be suitable for hybridization to the cDNA targets. In our hands *EcoR*I has produced consistently satisfactory results for human genomic DNA.
8. When beginning the technique, a range of different hybridization and wash temperatures should be tested to determine the optimal sensitivity and specificity for the specific cDNA microarrays used. In our hands (10) we have found that hybridization at 37ºC and wash at 55ºC allows sufficient sensitivity for detection of high copy number gains and amplifications. We have observed that 65ºC washes reduced signal intensity on our microarrays. Too low a wash temperature will result in non-specific binding (too many yellow signals). We recommend that these tests be performed using differentially labeled DNAs from different samples to ensure optimization of the technique specificity.
9. To date, the protocols for array fabrication have not yet been standardized. The published works specify target DNA concentrations of 400-1000 μg/mL hand-spotted on glass slides coated with poly-L-lysine (11), or 2 μg/μL target DNA on

aminopropyltrimethoxy silane-coated slides (12). It is important to note that both the concentration as well as the slide preparation is likely to change as automation procedures with robotic arrayers emerge.

10. The protocol specified herein for array CGH assumes a maximum gridded feature area that can be covered with a 22 x 20 mm^2 glass coverslip. In addition, it is assumed that the target DNA are denatured during array fabrication (12). Otherwise, a microarray denaturation step of 2 minutes in 70% formamide/4X SSC (11) must be included prior to probe hybridization.
11. The final probe length depends on the DNase I concentration. For CGH, the suitable length for hybridization ranges from 500-2,000 base pairs. Initially, stock solutions of $1x10^{-4}$ U/μL, prepared fresh in DNAse I dilution buffer, may be used to obtain the final concentration of $5x10^{-5}$ U/μL. However, this should be adjusted as necessary to obtain optimal fragment length.
12. Approximately 0.05 – 0.1 volume of each labeling mixture is loaded onto the gel with DNA stain (eg. ethidium bromide). Assessment of labeling by agarose gel is recommended as it can aid in troubleshooting array CGH results.

5. References

1. Kallioniemi, A., Kallioniemi, O. P., Sudar, D., Rutovitz, D., Gray, J. W., Waldman, F., Pinkel, D. (1992) Comparative genomic hybridization for molecular cytogenetic analysis of solid tumors. Science 258: 818-821
2. Kallioniemi, O. P., Kallioniemi, A., Sudar, D., Rutovitz, D., Gray, J. W., Waldman, F., Pinkel, D. (1993) Comparative genomic hybridization: a rapid new method for detecting and mapping DNA amplification in tumors. Seminars in Cancer Biology 4: 41-46.
3. Forozan, F., Karhu, R., Kononen, J., Kallioniemi, A., Kallioniemi, O. P. (1997) Genome screening by comparative genomic hybridization. Trends in Genetics 13: 405-409.
4. James, L. A. (1999) Comparative genomic hybridization as a tool in tumour cytogenetics. Journal of Pathology 187: 385-395.
5. Parente, F., Gaudray, P., Carle, G. F., Turc-Carel, C. (1997) Experimental assessment of the detection limit of genomic amplification by comparative genomic hybridization CGH. Cytogenetics & Cell Genetics 78: 65-68
6. Lennon, G., Auffray, C., Polymeropoulos, M., Soares, M. B. (1996) The I.M.A.G.E. Consortium: an integrated molecular analysis of genomes and their expression. Genomics 33: 151-152.
7. Schena, M., Shalon, D., Davis, R. W., Brown, P. O. (1995) Quantitative monitoring of gene expression patterns with a complementary DNA microarray. Science 270: 467-470.
8. DeRisi, J., Penland, L., Brown, P. O., Bittner, M. L., Meltzer, P. S., Ray, M., Chen, Y., Su, Y. A., Trent, J. M. (1996) Use of a cDNA microarray to analyse gene expression patterns in human cancer. Nature Genetics 14: 457-460.
9. Pollack, J. R., Perou, C. M., Alizadeh, A. A., Eisen, M. B., Pergamenschikov, A., Williams, C. F., Jeffrey, S. S., Botstein, D., Brown, P. O. (1999) Genome-wide analysis of DNA copy-number changes using cDNA microarrays. Nature Genetics 23: 41-46.
10. Beheshti, B., Braude, I., Marrano, P., Thorner, P., Zielenska, M., Squire, J. A. (2002) Chromosomal localization of DNA amplifications in neuroblastoma tumors using cDNA microarray comparative genomic hybridization. Neoplasia (submitted)
11. Solinas-Toldo, S., Lampel, S., Stilgenbauer, S., Nickolenko, J., Benner, A., Dohner, H., Cremer, T., Lichter, P. (1997) Matrix-based comparative genomic hybridization: biochips to screen for genomic imbalances. Genes, Chromosomes, & Cancer 20: 399-407.
12. Pinkel, D., Segraves, R., Sudar, D., Clark, S., Poole, I., Kowbel, D., Collins, C., Kuo, W. L., Chen, C., Zhai, Y., Dairkee, S. H., Ljung, B. M., Gray, J. W., Albertson, D. G. (1998) High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays. Nature Genetics 20: 207-211.
13. Heiskanen, M. A., Bittner, M. L., Chen, Y., Khan, J., Adler, K. E., Trent, J. M., Meltzer, P. S. (2000) Detection of gene amplification by genomic hybridization to cDNA microarrays. Cancer Research 60: 799-802.

14. Raap, A. K., van de Corput, M. P., Vervenne, R. A., van Gijlswijk, R. P., Tanke, H. J., Wiegant, J. (1995) Ultra-sensitive FISH using peroxidase-mediated deposition of biotin- or fluorochrome tyramides. Human Molecular Genetics 4: 529-534
15. Jones, T. A., Flomen, R. H., Senger, G., Nizetic, D., Sheer, D. (2000) The homeobox gene MEIS1 is amplified in IMR-32 and highly expressed in other neuroblastoma cell lines. European Journal of Cancer 36: 2368-2374.
16. Spieker, N., van Sluis, P., Beitsma, M., Boon, K., van Schaik, B. D., van Kampen, A. H., Caron, H., Versteeg, R. (2001) The MEIS1 oncogene is highly expressed in neuroblastoma and amplified in cell line IMR32. Genomics 71: 214-221.
17. Pei, J., Beheshti, B., Squire, J. A., Zielenska, M. (2001) cDNA array CGH analysis of pediatric osteosarcomas identifies regions of amplification on chromosome 17. Cancer Research in preparation
18. Halgren, R. G., Fielden, M. R., Fong, C. J., Zacharewski, T. R. (2001) Assessment of clone identity and sequence fidelity for 1189 IMAGE cDNA clones. Nucleic Acids Research 29: 582-588.
19. Albertson, D. G., Ylstra, B., Segraves, R., Collins, C., Dairkee, S. H., Kowbel, D., Kuo, W. L., Gray, J. W., Pinkel, D. (2000) Quantitative mapping of amplicon structure by array CGH identifies CYP24 as a candidate oncogene. Nature Genetics 25: 144-146.
20. Bruder, C. E., Hirvela, C., Tapia-Paez, I., Fransson, I., Segraves, R., Hamilton, G., Zhang, X. X., Evans, D. G., Wallace, A. J., Baser, M. E., Zucman-Rossi, J., Hergersberg, M., Boltshauser, E., Papi, L., Rouleau, G. A., Poptodorov, G., Jordanova, A., Rask-Andersen, H., Kluwe, L., Mautner, V., Sainio, M., Hung, G., Mathiesen, T., Moller, C., Pulst, S. M., Harder, H., Heiberg, A., Honda, M., Niimura, M., Sahlen, S., Blennow, E., Albertson, D. G., Pinkel, D., Dumanski, J. P. (2001) High resolution deletion analysis of constitutional DNA from neurofibromatosis type 2 (NF2) patients using microarray-CGH. Human Molecular Genetics 10: 271-282.
21. Daigo, Y., Chin, S. F., Gorringe, K. L., Bobrow, L. G., Ponder, B. A., Pharoah, P. D., Caldas, C. (2001) Degenerate oligonucleotide primed-polymerase chain reaction-based array comparative genomic hybridization for extensive amplicon profiling of breast cancers : a new approach for the molecular analysis of paraffin-embedded cancer tissue. American Journal of Pathology 158: 1623-1631.
22. Weber, T., Weber, R. G., Kaulich, K., Actor, B., Meyer-Puttlitz, B., Lampel, S., Buschges, R., Weigel, R., Deckert-Schluter, M., Schmiedek, P., Reifenberger, G., Lichter, P. (2000) Characteristic chromosomal imbalances in primary central nervous system lymphomas of the diffuse large B-cell type. Brain Pathology 10: 73-84
23. Geschwind, D. H., Gregg, J., Boone, K., Karrim, J., Pawlikowska-Haddal, A., Rao, E., Ellison, J., Ciccodicola, A., D'Urso, M., Woods, R., Rappold, G. A., Swerdloff, R., Nelson, S. F. (1998) Klinefelter's syndrome as a model of anomalous cerebral laterality: testing gene dosage in the X chromosome pseudoautosomal region using a DNA microarray. Developmental Genetics 23: 215-229

24. Hui, A. B., Lo, K. W., Yin, X. L., Poon, W. S., Ng, H. K. (2001) Detection of multiple gene amplifications in glioblastoma multiforme using array-based comparative genomic hybridization. Laboratory Investigation 81: 717-723.
25. Cheung, V. G., Nowak, N., Jang, W., Kirsch, I. R., Zhao, S., Chen, X. N., Furey, T. S., Kim, U. J., Kuo, W. L., Olivier, M., Conroy, J., Kasprzyk, A., Massa, H., Yonescu, R., Sait, S., Thoreen, C., Snijders, A., Lemyre, E., Bailey, J. A., Bruzel, A., Burrill, W. D., Clegg, S. M., Collins, S., Dhami, P., Friedman, C., Han, C. S., Herrick, S., Lee, J., Ligon, A. H., Lowry, S., Morley, M., Narasimhan, S., Osoegawa, K., Peng, Z., Plajzer-Frick, I., Quade, B. J., Scott, D., Sirotkin, K., Thorpe, A. A., Gray, J. W., Hudson, J., Pinkel, D., Ried, T., Rowen, L., Shen-Ong, G. L., Strausberg, R. L., Birney, E., Callen, D. F., Cheng, J. F., Cox, D. R., Doggett, N. A., Carter, N. P., Eichler, E. E., Haussler, D., Korenberg, J. R., Morton, C. C., Albertson, D., Schuler, G., de Jong, P. J., Trask, B. J. (2001) Integration of cytogenetic landmarks into the draft sequence of the human genome. Nature 409: 953-958.
26. Telenius, H., Carter, N. P., Bebb, C. E., Nordenskjold, M., Ponder, B. A., Tunnacliffe, A. (1992) Degenerate oligonucleotide-primed PCR: general amplification of target DNA by a single degenerate primer. Genomics 13: 718-725
27. Cheng, S., Fockler, C., Barnes, W. M., Higuchi, R. (1994) Effective amplification of long targets from cloned inserts and human genomic DNA. Proceedings of the National Academy of Sciences USA 91: 5695-5699.
28. Lichter, P., Joos, S., Bentz, M., Lampel, S. (2000) Comparative genomic hybridization: uses and limitations. Seminars in Hematology 37: 348-357.
29. Wilgenbus, K. K., Lichter, P. (1999) DNA chip technology ante portas. Journal of Molecular Medicine 77: 761-768
30. King, W., Proffitt, J., Morrison, L., Piper, J., Lane, D., Seelig, S. (2000) The role of fluorescence in situ hybridization technologies in molecular diagnostics and disease management. Molecular Diagnosis 5: 309-319.
31. Cheung, V. G., Morley, M., Aguilar, F., Massimi, A., Kucherlapati, R., Childs, G. (1999) Making and reading microarrays. Nature Genetics 21: 15-19.
32. Eisen, M. B., Spellman, P. T., Brown, P. O., Botstein, D. (1998) Cluster analysis and display of genome-wide expression patterns. Proceedings of the National Academy of Sciences USA 95: 14863-14868.
33. Brazma, A., Vilo, J. (2000) Gene expression data analysis. FEBS Letters 480: 17-24.
34. Kerr, M. K., Martin, M., Churchill, G. A. (2000) Analysis of variance for gene expression microarray data. Journal of Computational Biology 7: 819-837
35. Sherlock, G., Hernandez-Boussard, T., Kasarskis, A., Binkley, G., Matese, J. C., Dwight, S. S., Kaloper, M., Weng, S., Jin, H., Ball, C. A., Eisen, M. B., Spellman, P. T., Brown, P. O., Botstein, D., Cherry, J. M. (2001) The Stanford Microarray Database. Nucleic Acids Research 29: 152-155.
36. Phimister, B. (1999) Going Global (Editorial). Nature Genetics 21: 1
37. Alon, U., Barkai, N., Notterman, D. A., Gish, K., Ybarra, S., Mack, D., Levine, A. J. (1999) Broad patterns of gene expression revealed by clustering analysis of tumor and normal colon tissues probed by oligonucleotide arrays. Proceedings of the National Academy of Sciences USA 96: 6745-6750.

6. Figure Legends

Figure 1. Schematic depiction of the utility of cDNA microarrays in expression and CGH analyses. cDNA microarrays are screened with labeled probes derived from RNA and/or DNA of normal (Cy5) and tumour (Cy3) tissue. Analysis of the red:green signal intensity ratios indicate the level of A) gene expression or B) gene dosage change, respectively. Analyses may require datamining techniques for optimal interpretation of the results. A) Two-dimensional hierarchical clustering (32,37) is applied to the results to identify patterns of gene expression and establish clinical correlates. B) *in silico* cDNA chromosome localisation and arrangement into sequential order allows the results of cDNA array CGH to be depicted as an ideogram-type plot across the genome, facilitating identification of regions of gene dosage change.

Figure 2. Normalized cDNA array CGH of neuroblastoma patients identified three tumour genotypes: A) No high copy gains or amplification of genomic DNA; B) *MYCN* amplification as the sole genomic copy number imbalance; and C) *MYCN* amplification with previously undetected co-amplified 2p24 genes and high copy number gains of mitochondrial DNA sequences and numerous other genes, suggesting an underlying genetic instability. This third clinical genotype was not previously described, as these regions are not resolvable by chromosome CGH (10).

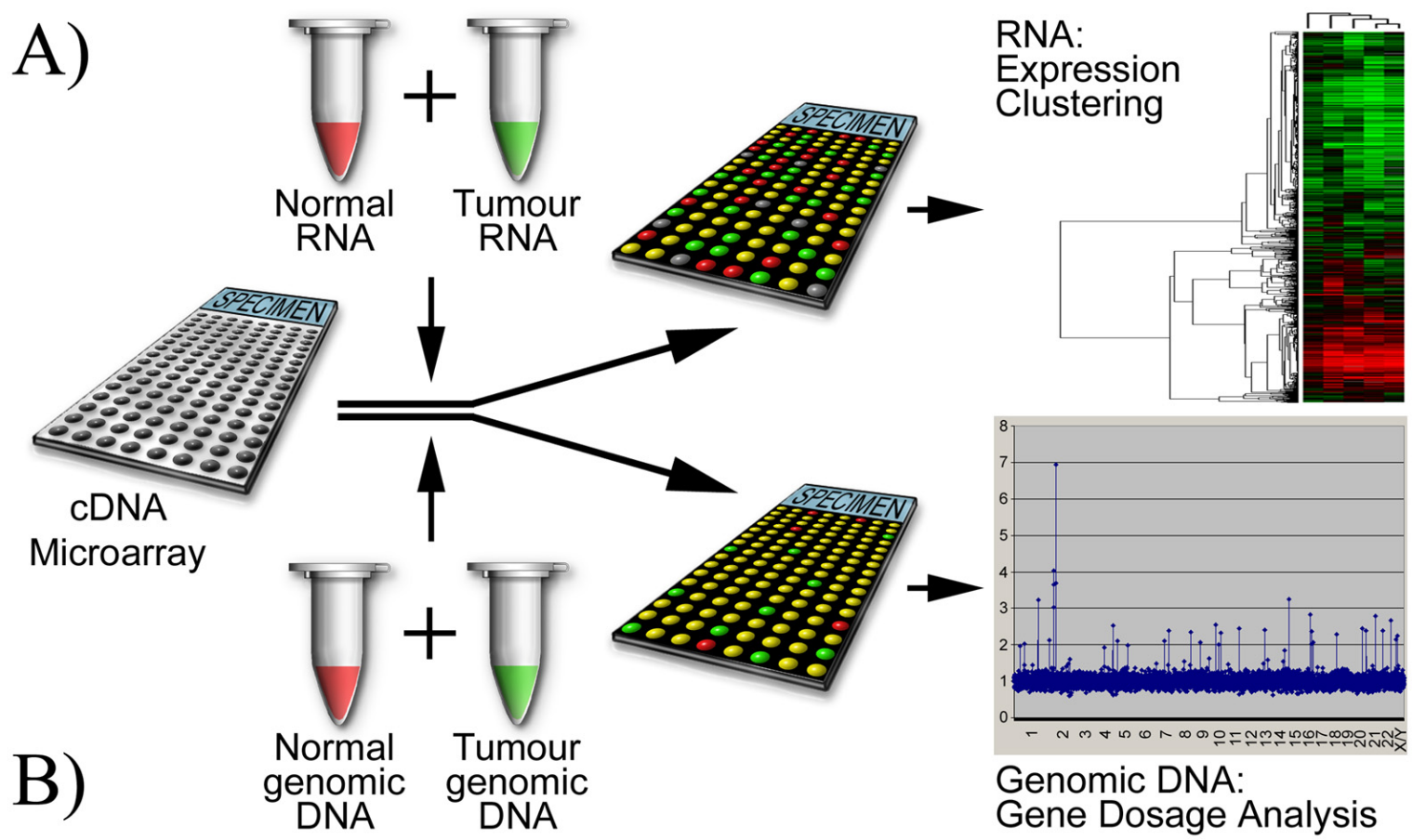
Figure 3. High resolution detection of gene dosage changes on chromosome 17 using high-density cDNA array CGH. Chromosome CGH detected high copy gain of the chromosome region 17p-17q21 (vertical gray bar) in an osteosarcoma sample.

Corresponding normalized cDNA array CGH using genomic DNA from the same sample significantly resolved the boundaries of this gain to the region 17p12-17p11.2 (horizontal gray bar). Chromosome ideograms are constructed by *in silico* assignment of microarray cDNAs to chromosomes, then arranging cDNAs into sequential order along each chromosome (10).

Figure 4. Schematic representation of the array CGH technique for a focused analysis of copy number imbalances along a region of interest (eg. 8q21.1). A) A tiling path of genomic clones (eg. BACs, PACs, P1s, cosmids) is generated to cover the region. After extraction and purification, these genomic DNA targets are arrayed onto glass slides. B) Array CGH is performed by hybridizing labeled normal (Cy3) and tumour (Cy5) genomic DNA to the microarray, and detected using a microarray scanner. C) Each array spot, realigned *in silico* as a single contiguous map to correspond with the tiling path, can be analysed by fluorescence ratio to identify the regions of copy number changes. These results may be correlated with *in silico* techniques to identify candidate genes of interest.

Figure 5. Histogram showing the copy number of the genomic clones comprising a 7 Mb tiling path on chromosome 22q, represented from the centromeric (left) to the telomeric (right) direction. Each black bar represents an individual genomic clone. Chromosome X and Y control genomic clones are separated (gray bar) on the right of the histogram. A) Array CGH comparing normal male and female DNA shows expected single copy loss of chromosome X clones (arrows). B) Comparison of a male NF2 patient against normal female control delineates boundaries of heterozygous loss along the *NF2* locus

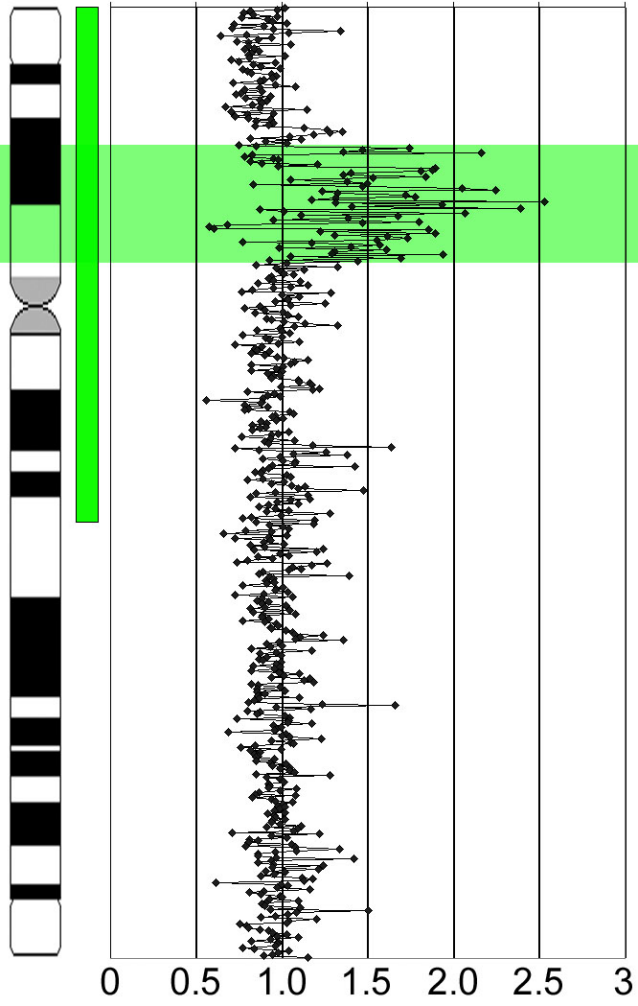
and surrounding region (stippled region). C) The detection of homozygous interstitial deletion (asterisk) within a region of single copy loss in a heterozygous female NF2 patient against a normal female control demonstrates the sensitivity and the resolution of array CGH. The accuracy of the technique is reflected by the deviation of the ratio from the expected values. Adapted from Bruder et al., 2001 (20) with permission.

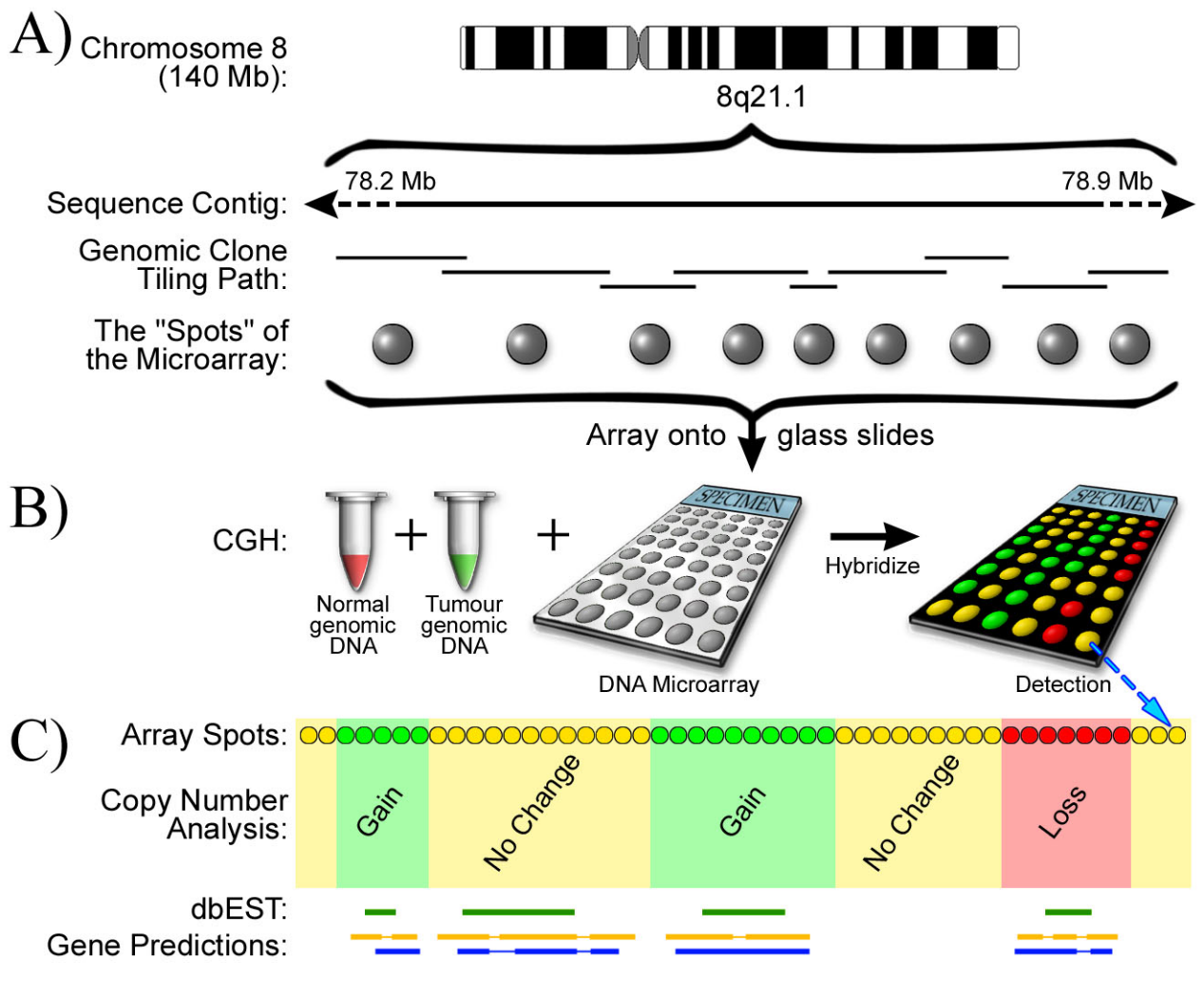

A)
Normal RNA
Tumour RNA
SPECIMEN
cDNA Microarray
SPECIMEN
RNA: Expression Clustering
SPECIMEN
Normal genomic DNA
Tumour genomic DNA
B)
Genomic DNA: Gene Dosage Analysis

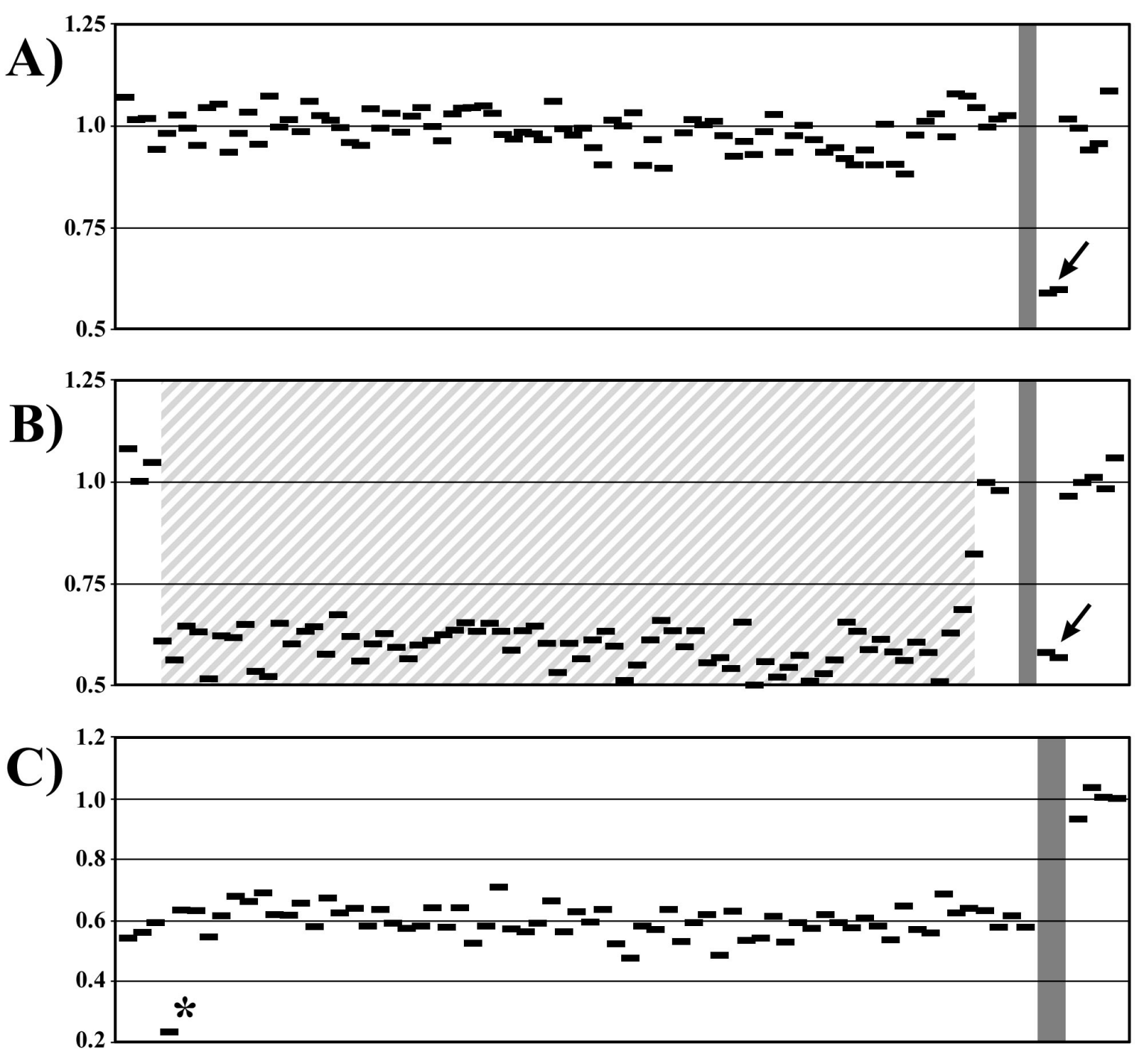
A
10
9
8
7
6
5
4
3
2
1
0
1p36.33
1p22.3
1q41
2q14.1
2p23.3
3q21.1
4q21.3
5q13.2
6p25.3
6q23.2
7q22.2
8q21.13
9q33.1
10q22.2
11q11
12p13.33
12q24.21
14q22.2
15q15.3
16p11.2
17q11.2
18q12.3
19q13.12
N/A
22q11
N/A
N/A
NF
NF
NF
NF
NF
NF
NF
NF
NF
X.q12

B
10
9
8
7
6
5
4
3
2
1
0
1p36.33
1p31.1
1q31.3
2p16.3
2q34
3p14.3
4p16.3
4
5q32
6p21.31
7p22.3
7q33
8q23.1
9q33.1
10q22.2
11p11.11
11
12q22
13q31.2
14q32.33
15q26.2
16q24.1
17q21.33
19p13.3
19q13.13
20pter-p
22q13.32
N/A
N/A
NF
NF
NF
NF
NF
NF
NF
X.q25

C
10
9
8
7
6
5
4
3
2
1
0
1p36.33
1p22.2
1q33.3
2p11.1
2q32.2
3q21.1
4q21.21
5q13.2
6p25.3
6q25.1
7q22.3
8q21.3
9q33.1
10q23.1
11q11
12p13.33
12q24.13
14q21.1
15q15.2
16p13.12
17q11.1
18q11.1
19q13.11
20q13.31
22q11.2
N/A
N/A
NF
NF
NF
NF
NF
NF
NF
NF
NF


0
0.5
1.0
1.5
2.0
2.5
3


A)
Chromosome 8 (140 Mb):
8q21.1
78.2 Mb
78.9 Mb
Sequence Contig:
Genomic Clone Tiling Path:
The "Spots" of the Microarray:
Array onto glass slides
B)
CGH:
Normal genomic DNA
Tumour genomic DNA
SPECIMEN
DNA Microarray
Hybridize
SPECIMEN
Detection
C)
Array Spots:
Copy Number Analysis:
Gain
No Change
Gain
No Change
Loss
dbEST:
Gene Predictions:


A)
1.25
1.0
0.75
0.5
B)
1.25
1.0
0.75
0.5
C)
1.2
1.0
0.8
0.6
0.4
0.2
*