

# Sensitivity issues in DNA array-based expression measurements and performance of nylon microarrays for small samples

François Bertucci<sup>2</sup>, Karine Bernard<sup>1</sup>, Béatrice Loriod<sup>1</sup>, Yi-Chung Chang<sup>3</sup>, Samuel Granjeaud<sup>1</sup>, Daniel Birnbaum<sup>2</sup>, Catherine Nguyen<sup>1</sup>, Konan Peck<sup>3</sup> and Bertrand R. Jordan<sup>1,+</sup>

TAGC, Institut de Cancérologie et d'Immunologie de Marseille: <sup>1</sup>CIML Luminy, 13288 Marseille Cedex 9, France, <sup>2</sup>Institut Paoli-Calmettes, Marseille, France and <sup>3</sup>Institute of Biomedical Sciences, Academia Sinica, Taipei, Taiwan

Received April 16, 1999; Revised and Accepted June 21, 1999

**DNA or oligonucleotide arrays are widely used for large-scale expression measurements, using various implementations: macroarrays in which DNA is spotted onto nylon membranes of relatively large dimensions (with radioactive detection) on the one hand; microarrays on glass slides and oligonucleotide chips, both used with fluorescent probes, on the other hand. Nylon microarrays with colourimetric detection have also been described recently. The small physical dimensions of miniaturized systems allow small hybridization volumes (2–100  $\mu$ l) and provide high probe concentrations, in contrast to macroarrays. We show, however, that actual sensitivity (defined as the amount of sample necessary for detection of a given mRNA species) is in fact similar for all these systems and that this is mostly due to the very different amounts of target material present on the respective arrays. We then demonstrate that the combination of nylon microarrays with <sup>33</sup>P-labelled radioactive probes provides 100-fold better sensitivity, making it possible to perform expression profiling experiments using submicrogram amounts of unamplified total RNA from small biological samples. This has important implications in basic and clinical research and makes this alternative approach particularly suitable for groups operating in an academic context.**

## INTRODUCTION

Large-scale measurement of gene expression using hybridization of complex probes prepared from total or mRNA [in this paper we keep to the usual convention whereby the (known) DNA fragments arrayed on nylon or glass are called 'targets' and the labelled sample 'probe'] to arrays of cDNA inserts or oligonucleotides is becoming a widely used technique both in the academic and in the corporate sectors (1,2). It is the logical and necessary counterpart of genome or EST sequencing and has the potential to provide a wealth of data in many areas of research. The method can provide quantitative measurement of

the expression levels of thousands of genes in different tissues or in normal versus pathological samples, with good reproducibility and freedom of artefacts in carefully controlled experiments. Sensitivity in these applications is often defined as the minimum detectable abundance level, e.g. 1/300 000, meaning that mRNA species present at this level (corresponding roughly to 1 molecule/cell) can be quantified. However, the amount of sample necessary to achieve this is an important issue in a number of situations where biological material is very limited, making microgram amounts of mRNA all but unobtainable. In this paper we compare the minimum amount of sample RNA necessary for measurement using: (i) nylon macroarrays (large arrays) with radioactive detection; (ii) nylon microarrays (small arrays) with colourimetric detection; (iii) glass microarrays with fluorescent detection; and (iv) oligonucleotide chips with fluorescent detection. We discuss the somewhat surprising results that emerge, i.e. roughly equivalent sensitivity (in this sense) in all cases, and show that this is due to the widely different amounts of target material in these four systems. We then show that nylon microarrays combined with radioactive labelling can provide sensitive expression measurement using very small quantities of sample, 100 times less than the other methods.

## RESULTS

### Basic principle of measurements and important parameters

These expression measurements all involve hybridization of complex probes at very low Cot (or Rot) values in conditions of large target excess for each individual sequence species. Equilibrium between bound and free probe is not reached during these experiments that, instead, display first order linear kinetics. Thus the signal observed on a given target after hybridization is proportional to the abundance of the corresponding sequence species in the complex probe [therefore to the abundance of the relevant mRNA in the tissue of origin (3)]. At the end of the experiment, hybridization is very far from the saturation level (i.e. 100% coverage of target by

<sup>+</sup>To whom correspondence should be addressed. Tel: +33 4 91 26 94 96; Fax: +33 4 91 26 94 30; Email jordan@ciml.univ-mrs.fr

**Table 1.** Comparative performance for nylon macro- and microarrays, glass microarrays and oligonucleotide chips, according to published reports and to the work reported in this paper

	Data from published reports (4,9,15,17)				Data from this study
	Nylon macroarrays	Nylon microarrays (colourimetric)	Glass microarrays	Oligonucleotide chips	Nylon microarrays (radioactive)
Targets	cDNA clones (colonies or PCR products)	cDNA clones (PCR products)	cDNA clones (PCR products)	20mer oligos synthesized <i>in situ</i>	cDNA clones (PCR products)
Support and format	50–2000 spots on an 8 × 12 cm <sup>2</sup> nylon membrane	9600 spots on a 2.7 × 1.8 cm <sup>2</sup> nylon membrane	50–6400 spots on a 1.8 × 1.8 cm <sup>2</sup> glass slide	64 000 oligonucleotides on a 1.28 × 1.28 cm <sup>2</sup> glass chip	200 spots on a 5 × 4 mm <sup>2</sup> nylon membrane
Sample amount	25 mg total RNA	1 mg mRNA	2 mg mRNA	10 mg mRNA	~0.1 mg total RNA
Detection	<sup>33</sup> P	Colourimetric	Fluorescence	Fluorescence	<sup>33</sup> P
Hybridization volume	40 ml	10 ml	2–10 ml	200 ml	100 ml
Image acquisition	Phosphor screen	Flat-bed scanner	Confocal scanning	Confocal scanning	Phosphor screen (high-resolution scanner)
Detection limit (mRNA abundance)	1/20 000	~1/20 000	1/100 000	1/300 000	1/10 000
Minimum sample amount for detection	25 × 10 <sup>6</sup> molecules	60 × 10 <sup>6</sup> molecules	20 × 10 <sup>6</sup> molecules	30 × 10 <sup>6</sup> molecules	0.2 × 10 <sup>6</sup> molecules

Sensitivity is expressed (last two lines) as the lower detection limit for mRNA abundance and also as the minimum number of molecules of a given sequence species that must be present in the starting sample (RNA) in order to measure a signal after array hybridization. Other important parameters (array size, amount and nature of sample) are summarized. For the first four columns, conditions are as given in the published papers (refs 4,9,15,17, respectively, for nylon macroarrays, glass microarrays, oligonucleotide chips and nylon microarrays); calculations are described in Materials and Methods. The last column shows corresponding figures from data reported in this paper.

probe) and the proportion of target molecules actually hybridized is quite small (<1%; see Materials and Methods).

### Definition of practical sensitivity and comparison in the current systems

Sensitivity in expression measurements is usually quoted as a fractional amount, e.g. 1/100 000, meaning that an mRNA species present at this abundance in the sample can be reliably detected. However, this figure is directly dependent on the amount of probe used and therefore on sample size: if the probe concentration in the hybridization mix is doubled, all signals (except possibly highly expressed genes that may be at the limit of the linear response region) will also double. In this discussion we will use a different parameter, the total number of molecules of one species that must be present in the sample in order to obtain a measurable signal on the corresponding target(s) in the experiment. This figure indicates how much starting material will be needed if genes expressed at a given relative level are to be detected in the experiment and is very important in practical terms.

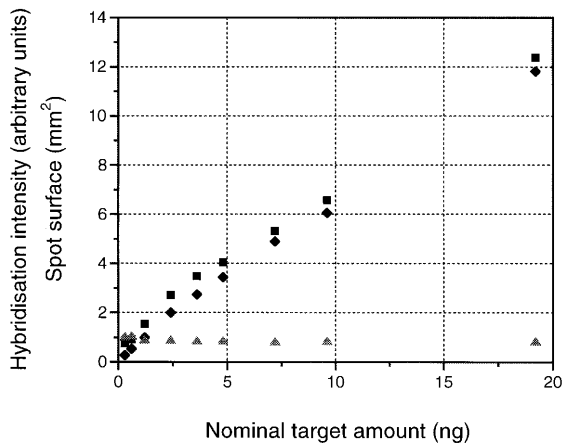
We have taken figures from published studies to compute this number (Materials and Methods) for the current implementations of expression measurement, macroarrays or high density membranes (3–8), glass microarrays (9–14), oligonucleotide chips (15,16) and nylon microarrays with colourimetric detection (17). The results are displayed in Table 1. In each case, calculations were based on a single paper, that giving most data on experimental conditions, i.e. Bernard *et al.* (4) for macroarrays, Chen *et al.* (17) for nylon microarrays, Schena (9) for glass microarrays and Lockhart *et al.* (15) for oligonucleotide chips. We have not taken into account the mRNA amplification methods used by some authors as: (i)

they are equally applicable to all four systems; and (ii) to our knowledge there is as yet no published proof that these procedures do not modify the relative abundances of individual sequence species. Corresponding figures for the work reported in this paper (see below) are indicated in the last column of Table 1.

The major conclusion from Table 1 is that sensitivity, as defined above, is very similar for the first four systems: 20–60 × 10<sup>6</sup> actual mRNA molecules of a given sequence species are required in the sample if one is to obtain a measurable hybridization intensity. This similarity between results with macroarrays on the one hand and miniaturized systems on the other is somewhat surprising as probe concentrations are much lower (by a factor of 10<sup>2</sup>–10<sup>4</sup>) in the macroarray system compared with the others. It is obviously interesting, and potentially very useful, to understand the causes of this discrepancy.

### Why are nylon macroarrays as sensitive as miniaturized systems?

*Amounts of target material are widely different and have a strong impact on the signal.* The most important difference between these systems is the amount of target material present in each spot or 'feature' (Materials and Methods). For colony macroarrays as in Bernard *et al.* (4), this has been estimated to be at least 20 ng of insert (3); for glass microarrays, the amount spotted is indicated (10) as 2.5 ng. Assuming insert fragments of 1 kb, this translates into 2 × 10<sup>10</sup> target molecules/spot for nylon macroarrays and 2.5 × 10<sup>9</sup> for glass microarrays; in both cases this assumes that all the material spotted has been bound to the support and is available for hybridization. For oligonucleotide chips, the number of identical oligonucleotides in



**Figure 1.** Variation of signal as a function of target amount. Nylon membranes ( $8 \times 12$  cm<sup>2</sup> macroarrays) on which various amounts of a reference *Arabidopsis* cDNA insert had been spotted (see text) were hybridized either with a complementary oligonucleotide probe under conditions of 10-fold probe excess (diamonds) or with a complex probe under conditions of very large target excess (squares). In both cases, the signal is approximately proportional to the amount deposited on the membrane. Spot surface as measured by the image analysis programme (triangles) does not vary significantly over this range.

one 'feature' is given as  $10^6$ – $10^7$  (15). Nylon microarrays are indicated by the authors to contain  $\sim 10$  ng DNA/spot (17).

Even though array hybridization occurs under conditions of large target excess, the amount of target material is still relevant to the results and we have reported, for colony arrays, that signals are in fact proportional to this amount (3). A simple experiment (Fig. 1) verifies this for macroarrays of PCR products spotted onto nylon membranes over a range extending from 0.3 to 19 ng of target DNA. A solution of insert material from a reference *Arabidopsis* clone (4) was deposited in increasing amounts onto a nylon membrane, using repeated spotting from the same solution with a solid pin device. The resulting spots (in octuplicate) received from 1 to 64 aliquots of target material, each corresponding to  $\sim 0.3$  ng ( $3 \times 10^8$  molecules) of insert. After processing of the membranes, one of them was hybridized with a 20mer oligonucleotide complementary to the insert under appropriate conditions (4) with a large excess of probe, to measure the amount of material bound to the membrane. As seen in Figure 1, the oligonucleotide signals increase almost linearly with the number of aliquots spotted. The quantification software used (18) also determines spot surfaces and shows that there is no detectable increase in apparent spot surface ( $\sim 0.8$  mm<sup>2</sup>) between single and multiple spottings.

An identical membrane was hybridized with a complex probe made from 25  $\mu$ g of total RNA to which had been added (before labelling) 1 ng of mRNA transcribed *in vitro* from the *Arabidopsis* clone. This corresponds to the reference procedure that we use to normalize results from different experiments (4) and mimics the detection of a species present at 0.2% abundance in the complex probe (assuming that 25  $\mu$ g of total RNA contain 500 ng of mRNA), under conditions of large target excess. As shown in Figure 1, the signals observed also follow the number of aliquots spotted with an essentially linear relationship. Thus signals in these complex probe hybridizations are indeed proportional not only to the abundance of individual species but also to the amount of target material.

**Rationale for target dependence of signal: target molecules are independent.** As indicated in Materials and Methods, estimation of the available surface and volume for DNA in a spot on a nylon membrane suggests that individual target DNA molecules are widely separated and are unlikely to interfere strongly with each other, in other words that they behave as independent entities in the hybridization reaction. Therefore, in a given spot, each target molecule has the same (small) probability of binding a complementary labelled probe molecule during the hybridization period. If the number of target molecules is doubled, individual probabilities remain the same but the total signal will also be doubled. Thus, and as indicated by the experiment shown in Figure 1, it is the large difference in target amounts between nylon and glass arrays (a difference that is likely to be much larger if binding efficiency and availability for hybridization are considered) that is responsible for the good sensitivity of the former, in spite of low probe concentrations.

In the case of nylon microarrays (17), the limited sensitivity of colourimetric detection offsets the advantage of high target amounts and small hybridization volumes and results in a figure similar to the three other methods. On the other hand, one major advantage of this approach is that results can be acquired using a very affordable flatbed scanner, instead of an expensive laser scanning system with confocal optics and photomultiplier detection.

**Other contributing factors.** The very small (2–10  $\mu$ l) volumes used for microarray hybridization (9,17) and the corresponding absence of macroscopic mixing of the hybridization solution may reduce the signal in these experiments because of local deprivation of probe in some areas, whereas macroarrays are hybridized in large volumes and under agitation, ensuring adequate mixing. For oligonucleotide chips, the 200  $\mu$ l volumes used (15) should allow mixing, but the number of target molecules is even smaller. The interference of bulky substituting groups added to DNA for detection by fluorescence may also lower the observed signals, while radioactive detection generally displays better intrinsic sensitivity.

In any case, these results, due to a combination of factors, indicate that nylon arrays, especially if they are somewhat miniaturized, can have a significant sensitivity advantage in terms of the amount of sample needed.

### Using nylon microarrays with radioactive detection

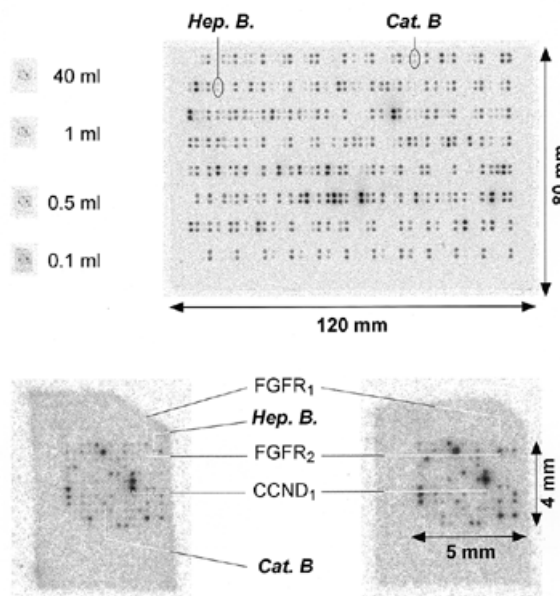
Given the good performance of nylon microarrays with a relatively insensitive colourimetric detection (17), we were interested in exploring the possibilities of this format using the more sensitive radioactive detection. Accordingly, small ( $4 \times 5$  mm<sup>2</sup>) nylon microarrays (300  $\mu$ m spacing) were constructed with PCR products from a set of IMAGE clones [estimated amount of target DNA per spot 10 ng (17)] corresponding to a series of cancer-related genes and used by us for expression profiling in breast cell lines and tumours (19). Several of these were hybridized with a <sup>33</sup>P-labelled complex probe prepared with RNA from the MDA-MB-134 breast adenocarcinoma cell line (American Type Culture Collection), in parallel with a conventional macroarray containing essentially the same set of sequences. This probe was prepared from 20  $\mu$ g of total RNA (4), corresponding to  $\sim 400$  ng of mRNA; this was then diluted into 42 ml of hybridization buffer. Forty millilitres of this mix

were used to hybridize one macroarray plus one microarray in a plastic box; three additional microarrays were hybridized with, respectively, 1, 0.5 and 0.1 ml of the same mix in small sealed plastic bags. After washing and exposure to an imaging plate (3 days), results were acquired by scanning at 100  $\mu\text{m}$  nominal resolution (Fuji Bas 1500).

As shown in Figure 2, strong hybridization was detected on all four microarrays, with no obvious difference related to hybridization volume. Resolution in the 100  $\mu\text{m}$  scan was not sufficient to quantify individual spots in the microarray, but a second exposure of the microarray that had been hybridized in 0.1 ml of solution followed by scanning at 25  $\mu\text{m}$  nominal resolution in a high resolution imaging plate system (Fuji Bas 5000) resolved the spots and showed results that can be compared with the data established using macroarrays. All four microarrays hybridized with the same probe in this experiment (Fig. 2) were analysed in this way; intensity values showed good reproducibility (>98% deviate by less than a factor of two), indicating that DNA spotting and hybridization kinetics were consistent for these different arrays. In Figure 2, the spots corresponding to cathepsin B and to the heparin-binding EGF-like growth factor are outlined. We have determined (19) that mRNA abundance for these two species in the MDA-MB-134 cell line is, respectively, 0.019 and 0.004% relative to total mRNA, using sets of control clones and spiking with the corresponding mRNA as previously described (4; see Materials and Methods). In the microarray hybridization shown, cathepsin B expression is easily detected; results with the other gene are somewhat marginal. We thus assign a conservative value of 1/10 000 for the detection limit in this experiment (Table 1, last column).

Quantitative analysis of the corresponding signals and evaluation of background level and fluctuation using previously described software and methods (3,4,18) gave the following results. Background values were determined for a series of spot-sized areas over the microarray and gave a value of 0.31 (arbitrary) units with a standard deviation of 0.07; total signals for cathepsin B and heparin-binding EGF-like growth factor were, respectively, 2.19 and 0.71. Thus the background-corrected values of 1.88 and 0.40 for these intensity values include an uncertainty due to background fluctuation of 0.07, i.e. <5% in the first case and close to 20% in the second. We can conclude that this experiment indeed measures specific mRNA abundance for species present at 0.019 and at 0.004% abundance (relative to total mRNA) with, respectively, good and acceptable precision. The amount of probe used for this microarray corresponds to 1/400 of the total sample, i.e. 0.05  $\mu\text{g}$  of total RNA or  $\sim 1$  ng of poly(A)<sup>+</sup> RNA. For cathepsin B, the amount present in the sample is 0.019% of this, i.e. 0.19 pg or  $\sim 2 \times 10^5$  molecules: reliable quantitative detection at this level represents a 100-fold improvement over published microarray (as well as oligonucleotide chip and macroarray) performance. We have checked (data not shown) that probe labelling can be performed without difficulty using our protocols (4) on submicrogram amounts of total RNA.

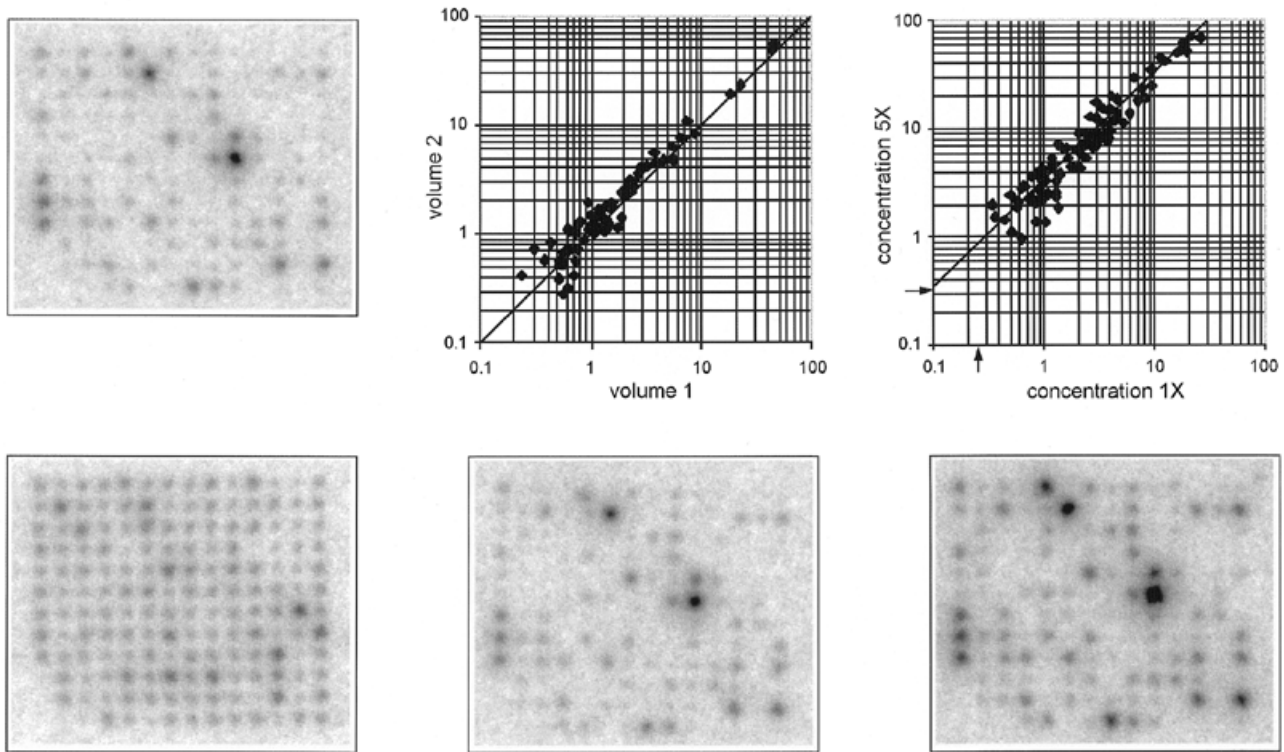
Figure 2 also shows a high resolution scan of an identical microarray hybridized at the same probe concentration with material from a different origin, the Kato III gastric carcinoma cell line (American Type Culture Collection), and demonstrates clear differential expression for a number of genes. For example, fibroblast growth factor receptor 1 (FGFR1) is clearly expressed in MDA-MB-134 but not in Kato III cells; cyclin D1 (CCND1) is strongly expressed in the first line but



**Figure 2.** Parallel hybridization of nylon macro- and microarrays. (Top) A macroarray and (left) four microarrays carrying essentially the same set of cDNAs corresponding to human genes potentially involved in cancer, hybridized with a <sup>33</sup>P-labelled complex probe prepared from a breast cancer cell line. After dilution of the probe in 42 ml of buffer, the macroarray was hybridized in 40 ml with one microarray; the other three were incubated with, respectively, 1, 0.5 and 0.1 ml of the same hybridization mix. After washing, the macroarray and the four microarrays were exposed together to a phosphor screen and scanned at 100  $\mu\text{m}$  nominal resolution in a Fuji BAS 1500 instrument. The 192 control clones used for signal normalization (Materials and Methods) are seen on the macroarray as a regular vertical grid. (Bottom left) The microarray hybridized in 0.1 ml was exposed to a new phosphor screen that was scanned at 25  $\mu\text{m}$  nominal resolution in a Fuji BAS 5000 instrument. Two genes expressed at relatively low levels are indicated on the macroarray (double spotting) and on the microarray. (Bottom right) An identical microarray, hybridized under the same conditions with a probe prepared from a stomach carcinoma cell line. Three differentially expressed genes are indicated on the two microarrays.

very weakly in the other; fibroblast growth factor receptor 2 (FGFR2), absent in the first case, is well expressed in the second. The results are in agreement with data from the literature (20–22) and confirm that these arrays can be used for very sensitive differential expression profiling.

An increase in probe concentration should provide stronger signals and a corresponding improvement in the detection limit expressed as fractional mRNA abundance, providing that the increase in background is limited. An indication that this is possible is presented in Figure 3, which shows results obtained under different conditions. It first highlights the similarity of images and data obtained with the same probe concentration but widely different hybridization volumes, then the effect of increasing probe concentration by a factor of five. It is apparent that signals increase strongly while background is only slightly elevated. In quantitative terms, background (determined as above) increases from  $0.26 \pm 0.08$  to  $0.34 \pm 0.11$  (arbitrary units), i.e. by  $\sim 30\%$ , while signals increase by an average factor of 3.7. Thus the detection limit is improved by an approximate factor of 3. Figure 3 also shows the results of hybridizing an array with an oligonucleotide corresponding to



**Figure 3.** Hybridization of nylon microarrays under various conditions. Four microarrays from the same batch were hybridized under different conditions and numerical results compared (log scale). (Left) Hybridization at 'low' complex probe concentration (top) and with a vector oligonucleotide to assay the amount of DNA in each spot (bottom). (Middle) Numerical comparison of results with different volumes (volume 1, 40 ml; volume 2, 0.5 ml) of the same probe at the same concentration (top) with the corresponding image (bottom). (Right) Numerical comparison of results with a 5-fold change in probe concentration and indication of the background levels (top; arrows); (bottom) corresponding image.

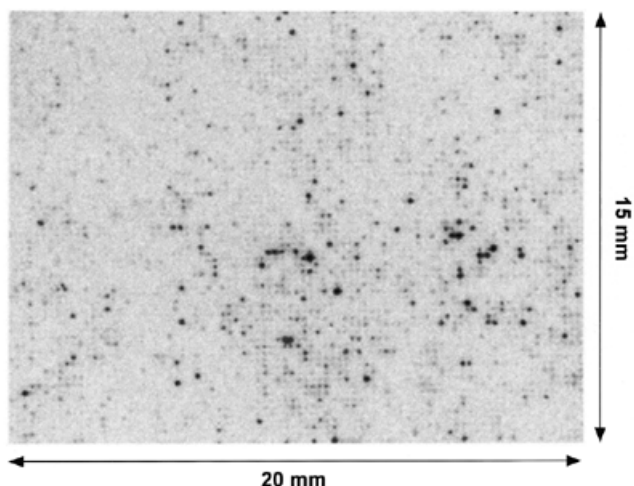
a polylinker sequence present in all of the PCR products spotted, in order to quantify the amount of target DNA present in each spot. This hybridization is done under excess probe conditions (4) in order to ensure the most reliable results. Quantification of these data shows an approximately Gaussian distribution with a mean value of 3.2 arbitrary units and a standard deviation of 1.5, plus a few more intense signals (which do not correspond to the two genes discussed above). These data can be used to correct the intensity values obtained with the complex probe, as previously discussed (4); it also indicates the need for better standardization of PCR product concentration and deposited volumes, with more sophisticated devices than the simple solid pin system (17) used to fabricate these arrays.

Higher density microarrays containing 9600 human cDNAs on a  $2.7 \times 1.8 \text{ cm}^2$  nylon membrane have been used with colourimetric detection (17). Radioactive sampling of these arrays would also be of interest, but the pitch of 225  $\mu\text{m}$  makes imaging plate detection marginal. However, real-time detectors primarily designed for *in situ* tissue hybridization can be used for this purpose; Figure 4 shows an image acquired with a Biospace MicroImager of such an array hybridized with a  $^{33}\text{P}$ -labelled complex probe prepared with 5  $\mu\text{g}$  of total RNA ( $\sim 100 \text{ ng}$  of mRNA) from the MDA-MB-134 cell line and used at low concentration (10 ml total hybridization volume). Resolution is indeed adequate and shows that such an instrument can assay  $^{33}\text{P}$  on this microscale; dynamic range, determined by compar-

ing counts for strong and weak spots, extends over three decades. Sensitivity, as assessed by the fraction of spots giving a detectable signal ( $\sim 50\%$ ), is limited in this experiment, but can be improved by hybridization at higher probe concentrations. And, of course, the amount of sample material used is very small. We believe that this format may provide a good approach to high density, high sensitivity expression profiling.

## DISCUSSION

In the recent trend towards miniaturization of high throughput expression measurement, nylon arrays have been under-emphasized, largely because of the apparently unavoidable background fluorescence of the material that severely limits sensitivity using fluorescence detection methods. However, the high binding capacity of nylon membranes does allow good results with colourimetric (enzymatic) detection, with very high spot density (9600 on a  $2.7 \times 1.8 \text{ cm}^2$  array) and minimal instrument requirements (17). The results presented in this paper show that nylon microarrays, when used with radioactive detection, allow expression profiling using very small samples of cells or tissues. We have shown quantification of expression with a total number of specific sample molecules of  $\sim 2 \times 10^5$ , i.e. 100 times less than all other methods. This translates, for detection down to an abundance level of 1/10 000, to a total of only  $2 \times 10^9$  mRNA molecules in the sample, i.e. 100 ng of total RNA ( $\sim 10^4$  cells). Such results open the possibility



**Figure 4.** High resolution radioactive probing of a nylon microarray. A  $2.7 \times 1.8$  cm<sup>2</sup> nylon microarray (17) with 225  $\mu$ m pitch, hybridized with a <sup>33</sup>P-labelled complex probe prepared from MDA-MB-134 cell line RNA. Image acquired using a Biospace Micro Imager instrument (exposure time 3 days). The area shown is actually  $2.0 \times 1.5$  cm<sup>2</sup> (corresponding to the 'window' of the particular instrument used) and contains ~6000 cDNAs.

of using unamplified material from flow-sorted cells, dissected tumours or even biopsies, a very important element in a number of basic and clinical research situations.

The aspect of sensitivity we have emphasized here (minimum number of sample molecules for detection) is complementary to a more commonly discussed parameter, the minimum detectable relative abundance limit. Glass microarrays and oligonucleotide chips have proven capable of detecting mRNA species at 1/300 000 relative abundance (~1 molecule/cell), a figure that has not been reported for nylon arrays. Thus, for applications requiring detection of very rare species and when sample is not limited, these approaches may have definite advantages.

A few practical issues also require comment. Nylon microarrays can be manufactured with commercially available robots, especially those that do not use 'split pin' systems for sample deposition and rely instead on piezoelectric devices or solid pins (23); the relative fragility of this material is not a problem given the small size of the arrays. They require relatively large amounts of PCR products; however, these need no elaborate purification before application to the small nylon membranes (17) whereas this is necessary in glass microarray applications. Nylon arrays can also, again in contrast to glass microarrays or oligonucleotide chips, be reused several times after probe stripping. The high resolution imaging plate systems or direct radioactivity detection instruments are relatively expensive and not generally available in research laboratories, although they are in the same price range as the laser scanners with confocal optics used for glass microarrays or oligonucleotide chips.

The limited resolution of current radioactivity detection systems makes miniaturization beyond the scale demonstrated in this paper (225  $\mu$ m pitch) unlikely. For very large-scale expression profiling of tens of thousands of genes, present (and future) technologies using optical detection probably provide a better approach. However, for investigation of situations where

a smaller number of genes need to be assayed or when sample material is scarce (flow-sorted cells, clinical samples, etc.) the alternative we describe provides very significant advantages while retaining high flexibility and blending well into present research laboratory practices.

## MATERIALS AND METHODS

### Membranes, hybridization and detection

Macroarrays were prepared by spotting solutions of PCR products from IMAGE cDNA clones onto Hybond N+ membranes using a commercial gridding robot (BioGrid; BioRobotics, Cambridge, UK) as described (3,19). Microarrays on the same support were prepared using the solid pin microspotting device developed by some of us at Academia Sinica (17). The set of clones used in these experiments can be accessed at <http://tagc.univ-mrs.fr/pub/Cancer/>. Complex probes were prepared from total RNA as previously described (4) with <sup>33</sup>P labelling. Hybridization was for 48 h at 68°C in the volumes indicated (in plastic boxes for large volumes, in small plastic bags for volumes below 1 ml). After washing, the arrays were exposed for 1–5 days on imaging plates that were scanned either in a Fuji BAS 1500 instrument (100 or 200  $\mu$ m nominal resolution) or a Fuji BAS 5000 machine (25  $\mu$ m nominal resolution). Acquisition of radioactive images at very high resolution from a nylon microarray containing 9600 human cDNA inserts (17) was performed with a Micro Imager, a high resolution digital autoradiography system displaying in real time the quantitative image of radioisotopes deposited on biological samples and providing 15  $\mu$ m spatial resolution with 5  $\mu$ m pixel size (Biospace Mesures, Paris, France). After image acquisition, the hybridization signals were quantified using the HD grid analyser software from Genomic Solutions (18). Signal processing was done using vector oligonucleotide data to correct for the relative amount of DNA present in each spot (for both macroarrays and microarrays); relative abundances were determined on the macroarray with reference to control clones (*Arabidopsis thaliana* cytochrome c554) detected with a known amount of homologous probe, as previously described (4).

### Calculation of parameters

*Fraction of target hybridized.* In previous complex probe hybridizations (3), we have measured the radioactive probe concentration in the hybridization buffer at the start and the end of the experiment. Using a whole thymus complex probe on an array of 1536 clones from an unnormalized thymus cDNA library, probe consumption measured in this way was ~20% of the initial amount. Thus a maximum of 20% of the probe is bound to the targets; this is an overestimate as part of the probe is loosely bound and will be removed by stringency washes. Relative amounts of probe and target, and average target coverage in this case, can then be estimated as follows. The probe is prepared by oligo(dT)-primed reverse transcription from 25  $\mu$ g of total RNA containing ~500 ng mRNA (4). The amount of labelled probe is thus <500 ng (as reverse transcription is not 100% efficient). The 20% of probe bound to the targets thus corresponds to <100 ng of material; this figure does not depend on the fraction of probe sequences represented in the array. On the other hand, the total amount of target material, using the figure of 20 ng DNA/spot (see below), is ~30 000 ng ( $20 \times 1536$ ). Thus, an overestimated value for average

probe coverage of targets at the end of hybridization is 100/30 000 (0.34%), the real figure being probably significantly lower.

**Calculation of sensitivity from published data.** For these calculations, we assume that mRNA represents 2% of total RNA for a typical mammalian cell, that the average size of an mRNA molecule is 1.7 kb and that the cDNA inserts present on macro- or microarrays are 1 kb long. With these figures, 1 pg of mRNA and 1 pg of (double-stranded) cDNA insert both correspond to  $\sim 10^6$  molecules.

For macroarrays, we use the figures quoted in Bernard *et al.* (4): detection of species expressed at a 1/20 000 level using 25  $\mu\text{g}$  of total RNA [500 ng of poly(A)<sup>+</sup> RNA]. Thus the minimum amount of sample for detection of a molecular species is equal to (500/20 000) ng, i.e. 25 pg or  $25 \times 10^6$  mRNA molecules. For nylon microarrays with colourimetric detection (17), the value is provided by the authors in the form  $60 \times 10^6$  mRNA molecules. For glass microarrays, the figures given in Schena (9) are 1/100 000 using 2  $\mu\text{g}$  of poly(A)<sup>+</sup> RNA, i.e. a minimum of 20 pg for detection. Finally, Lockhardt *et al.* (15) quote a minimum value of 30 pg for oligonucleotide chip measurements; this can also be calculated from the indicated sensitivity figure of 1/300 000 (1 molecule/mammalian cell) using 10  $\mu\text{g}$  of mRNA. The minimum amount of sample necessary for detection provides an indication of sensitivity that is independent of the total amount of RNA used to prepare the probe and thus allows a more straightforward comparison of different systems and experimental conditions.

**Amount of target in various arrays.** We determined previously that the amount of plasmid insert DNA in a typical bacterial colony on a nylon macroarray was 20–60 ng (3); experiments (data not shown) using spotted PCR products of known concentration and giving equivalent signals confirm this. A value of 20 ng corresponds to  $\sim 20 \times 10^9$  molecules of a 1 kb insert per spot. For glass microarrays, Schena *et al.* (10) indicate spotting 0.005  $\mu\text{l}$  of a 0.5 mg/ml solution, i.e. 2.5 ng ( $2.5 \times 10^9$  molecules/microarray spot, assuming 100% binding and 100% availability for hybridization, both figures being certainly overestimated). Other estimates from manufacturer's literature suggest that the amounts spotted are in fact a fraction of a nanogram. Finally, Lockhardt *et al.* (15) quote the number of identical oligonucleotides per 'feature' on Affymetrix glass chips as  $10^6$ – $10^7$ , with typically 20 different 20mers as targets for each gene to be assayed; assuming that each oligonucleotide species acts as an independent target, this translates into  $10^7$ – $10^8$  targets/gene. The 20 mismatched oligonucleotides included for background measurement (15) are not taken into account in this evaluation.

**Physical parameters for a DNA spot on a nylon membrane.** The nylon membrane (Hybond N+) used for macro- and microarrays has finite thickness ( $\sim 100 \mu\text{m}$ ) and an open pore sponge-like structure. Atomic force microscopy (AFM) of such a membrane (M. Marilley *et al.*, personal communication) shows a very uneven surface with  $\sim 1 \mu\text{m}$  peaks and troughs; the actual total surface of a  $4 \times 4 \mu\text{m}^2$  square was determined to be  $66 \mu\text{m}^2$ . Such a surface provides many sites for DNA binding.

In terms of volume, the total volume occupied by the target DNA can be computed as follows for macroarrays such as the ones used in Figures 1 and 2. Assuming that the 1 kb DNA target

segments are  $1 \mu\text{m}$  long with a diameter of  $2 \times 10^{-3} \mu\text{m}$  (20 Å), the individual volume occupied by one molecule is  $\sim 3 \times 10^{-6} \mu\text{m}^3$ , i.e.  $6 \times 10^4 \mu\text{m}^3$  for  $2 \times 10^{10}$  molecules (20 ng). Assuming that DNA binding occurs down to a depth of 25  $\mu\text{m}$  in the  $\sim 100 \mu\text{m}$  thick nylon membrane, the total volume of a spot of 0.8 mm diameter is of the order of  $10^7 \mu\text{m}^3$ . The fraction of this volume not occupied by the nylon fibres is estimated (from the AFM data) to be of the order of 50%; this corresponds to  $5 \times 10^6 \mu\text{m}^3$ , i.e. almost 100 times the total volume of the DNA molecules.

## ACKNOWLEDGEMENTS

We thank Monique Marilley, José Rocca-Serra and Jacques Barbet (UPRES 2059, Université de Provence, Marseille) for performing the AFM studies and calculations and Rémi Houllatte for critical discussion of the manuscript. We are very grateful to Vincent Tarazona (Raytest) for making a BAS 5000 instrument available to us and to Fabienne Tardieu (Biospace) for performing image acquisition on a Micro Imager. We thank Claude Mawas and Dominique Maraninchi for encouragement and support through the Institut Paoli-Calmettes. This work was supported by institutional grants to CIML from CNRS and INSERM and by specific grants to B.R.J. and D.B. from Association Française contre les Myopathies, Ligue Nationale Contre le Cancer and Association pour la Recherche sur le Cancer.

## REFERENCES

- Ramsay, G. (1998) DNA chips: state-of-the art. *Nature Biotechnol.*, **16**, 40–44.
- Jordan, B.R. (1998) Large-scale expression measurement by hybridization methods: from high-density membranes to 'DNA chips'. *J. Biochem. (Tokyo)*, **124**, 251–258.
- Nguyen, C., Rocha, D., Granjeaud, S., Baldit, M., Bernard, K., Naquet, P. and Jordan, B.R. (1995) Differential gene expression in the murine thymus assayed by quantitative hybridization of arrayed cDNA clones. *Genomics*, **29**, 207–215.
- Bernard, K., Auphan, N., Granjeaud, S., Victorero, G., Schmitt-Verhulst, A.M., Jordan, B.R. and Nguyen, C. (1996) Multiplex messenger assay: simultaneous, quantitative measurement of expression of many genes in the context of T cell activation. *Nucleic Acids Res.*, **24**, 1435–1443.
- Lennon, G.G. and Lehrach, H. (1991) Hybridization analyses of arrayed cDNA libraries. *Trends Genet.*, **7**, 314–317.
- Gress, T.M., Hoheisel, J.D., Lennon, G.G., Zehetner, G. and Lehrach, H. (1992) Hybridization fingerprinting of high-density cDNA-library arrays with cDNA pools derived from whole tissues. *Mamm. Genome*, **3**, 609–619.
- Zhao, N., Hashida, H., Takahashi, N., Misumi, Y. and Sakaki, Y. (1995) High-density cDNA filter analysis: a novel approach for large-scale, quantitative analysis of gene expression. *Gene*, **156**, 207–213.
- Pietu, G., Alibert, O., Guichard, V., Lamy, B., Bois, F., Leroy, E., Mariage-Samson, R., Houllatte, R., Soularue, P. and Auffray, C. (1996) Novel gene transcripts preferentially expressed in human muscles revealed by quantitative hybridization of a high density cDNA array. *Genome Res.*, **6**, 492–503.
- Schena, M. (1996) Genome analysis with gene expression microarrays. *Bioessays*, **18**, 427–431.
- Schena, M., Shalon, D., Davis, R.W. and Brown, P.O. (1995) Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science*, **270**, 467–470.
- Chu, S., DeRisi, J., Eisen, M., Mulholland, J., Botstein, D., Brown, P.O. and Herskowitz, I. (1998) The transcriptional program of sporulation in budding yeast. *Science*, **282**, 699–705.
- Marton, M.J., DeRisi, J.L., Bennett, H.A., Iyer, V.R., Meyer, M.R., Roberts, C.J., Stoughton, R., Burchard, J., Slade, D., Dai, H., Bassett, D.E.Jr, Hartwell, L.H., Brown, P.O. and Friend, S.H. (1998) Drug target validation and identification of secondary drug target effects using DNA microarrays. *Nature Med.*, **4**, 1293–1301.

13. Iyer, V.R., Eisen, M.B., Ross, D.T., Schuler, G., Moore, T., Lee, J.C.F., Trent, J.M., Staudt, L.M., Hudson, J.Jr, Boguski, M.S., Lashkari, D., Shalon, D., Botstein, D. and Brown, P.O. (1999) The transcriptional program in the response of human fibroblasts to serum. *Science*, **283**, 83–87.
14. Derisi, J.L., Iyer, V.R. and Brown, P.O. (1997) Exploring the metabolic and genetic control of gene expression on a genomic scale. *Science*, **278**, 680–686.
15. Lockhart, D.J., Dong, H., Byrne, M.C., Follettie, M.T., Gallo, M.V., Chee, M.S., Mittmann, M., Wang, C., Kobayashi, M., Horton, H. and Brown, E.L. (1996) Expression monitoring by hybridization to high-density oligonucleotide arrays. *Nature Biotechnol.*, **14**, 1675–1680.
16. Wodicka, L., Dong, H., Mittmann, M., Ho, M.-H. and Lockhart, D.J. (1997) Genome-wide expression monitoring in *Saccharomyces cerevisiae*. *Nature Biotechnol.*, **15**, 1359–1367.
17. Chen, J.J.W., Wu, R., Yang, P.C., Huang, J.Y., Sher, Y.P., Han, M.H., Kao, W.C., Lee, P.J., Chiu, T.F., Chang, F., Chu, Y.W., Wu, C.W. and Peck, K. (1998) Profiling expression patterns and isolating differentially expressed genes by cDNA microarray system with colorimetry detection. *Genomics*, **51**, 313–324.
18. Grangeaud, S., Nguyen, C., Rocha, D., Luton, R. and Jordan, B.R. (1996) From hybridization image to numerical values: a practical, high throughput quantification system for high density filter hybridizations. *Genet. Anal. Biomol. Eng.*, **12**, 151–162.
19. Bertucci, F., Van Hulst, S., Bernard, K., Llorca, B., Grangeaud, S., Tagett, R., Starkey, M., Nguyen, C., Jordan, B.R. and Birnbaum, D. (1999) Expression scanning of an array of growth control genes in human tumor cell lines. *Oncogene*, **18**, 3905–3912.
20. Penault-Llorca, F., Bertucci, F., Adelaïde, J., Parc, P., Coulier, F., Jacquemier, J., Birnbaum, D. and Delapeyrière, O. (1995) Expression of FGF and FGF receptor genes in human breast cancer. *Int. J. Cancer*, **61**, 170–176.
21. Sweeney, K.J., Swarbrick, A., Sutherland, R.L. and Musgrove, E.A. (1998) Lack of relationship between CDK activity and G1 cyclin expression in breast cancer cells. *Oncogene*, **16**, 2865–2878.
22. Kishi, T., Yoshida, T. and Terada, M. (1994) A soluble form of K-sam/FGFR2 protein in the culture medium of human gastric cancer cells. *Biochem. Biophys. Res. Commun.*, **202**, 1387–1394.
23. Bowtell, D.D. (1999) Options available—from start to finish—for obtaining expression data by microarray. *Nature Genet.*, **21** (Suppl. 1), 25–32.