

Specificity assessment from fractionation experiments (SAFE): a novel method to evaluate microarray probe specificity based on hybridisation stringencies

Alexei L. Drobyshev, Christine Machka, Marion Horsch, Matthias Seltmann, Volkmar Liebscher¹, Martin Hrabé de Angelis and Johannes Beckers*

Institute of Experimental Genetics and ¹Institute of Biomathematics and Biometry, GSF–National Research Centre for Environment and Health, Ingolstaedter Landstrasse 1, 85764 Neuherberg, Germany

Received August 26, 2002; Revised October 11, 2002; Accepted November 1, 2002

ABSTRACT

The cDNA-chip technology is a highly versatile tool for the comprehensive analysis of gene expression at the transcript level. Although it has been applied successfully in expression profiling projects, there is an ongoing dispute concerning the quality of such expression data. The latter critically depends on the specificity of hybridisation. SAFE (specificity assessment from fractionation experiments) is a novel method to discriminate between non-specific cross-hybridisation and specific signals. We applied *in situ* fractionation of hybridised target on DNA-chips by means of repeated washes with increasing stringencies. Different fractions of hybridised target are washed off at defined stringencies and the collected fluorescence intensity data at each step comprise the fractionation curve. Based on characteristic features of the fractionation curve, unreliable data can be filtered and eliminated from subsequent analyses. The approach described here provides a novel experimental tool to identify probes that produce specific hybridisation signals in DNA-chip expression profiling approaches. The iterative use of the SAFE procedure will result in increasingly reliable sets of probes for microarray experiments and significantly improve the overall efficiency and reliability of RNA expression profiling data from DNA-chip experiments.

INTRODUCTION

Arrays of immobilised cDNAs or oligonucleotides are emerging as a universal and versatile tool for the functional analysis of RNA expression profiles (1–5). Gene expression profiling using the DNA-chip technology has proven useful and powerful for the analysis of molecular pathways in the molecular network of the cell. A comprehensive transcriptome

analysis in a compendium of yeast mutants has led to the identification of new gene functions and co-regulated *syn*-expression groups of genes (6). In *Drosophila*, the DNA-chip technology has been used to study molecular pathways during metamorphosis (7), and in human cancer research expression profiling has provided new insights into pathogenesis and in the classification of tumours (8–10) and inflammatory diseases (11).

Comprehensive genome wide expression profiling has been suggested to be one of the tools in the worldwide effort to annotate the mammalian genome with biological functions (12,13). Whereas the current knowledge of gene function is usually limited to single pathways or a small set of target genes, transcription profiling of mouse mutant lines (their organs or derived cell lines) or of mice challenged by infectious disease allows a comprehensive analysis of interactions in global regulatory networks. Several recent reports have successfully used DNA microarray technologies for transcriptome analysis in mice. For example, the transcriptional response to ageing in the mouse brain has significant similarities to that in human neurodegenerative disorders, such as Alzheimer's disease (14,15). The differential gene expression in several brain regions and the response to seizure has also been analysed and provided evidence that particular differences in gene expression may account for distinct phenotypes in mouse inbred strains (16). These and further reports (17–19) have provided the proof-of-principle that despite the complexity of mammalian organs, expression profiling is a useful tool to identify pathways associated with particular biological processes in the mouse model system.

The reliability of expression profile data obtained in DNA-chip experiments is a major concern for the exact appraisal of differential gene expression (20). The repetition of experiments (21) and replicates of clones in an array (21,22) are standard procedures often used to support the reliability of expression data. However, such procedures cannot exclude the generation of false data. Artefacts can be due to particular probe sequences and structures that cause cross-hybridisation or the biased labelling with fluorescent dyes and the label itself. Such false data may therefore be highly reproducible.

*To whom correspondence should be addressed. Tel: +49 89 3187 3513; Fax: +49 89 3187 3500; Email: beckers@gsf.de

Another approach is the use of several different sequences corresponding to the same mRNA. The number of such probes for one specific gene may be as high as 40 in commercial microarrays (23). This strategy requires a high number of specific oligonucleotides per gene, is expensive and relies on the presumption that the majority of probes for each gene produce specific hybridisation, which is not valid *a priori*. Here we describe methods to verify the quality of each individual probe immobilised on an array in relation to the target RNA used for hybridisation. Although we apply this technology to tissues from mouse lines, this procedure may equally be applied to biological material from other sources.

It has been shown that melting of double-stranded DNA in solution can be described as a melting curve with sigmoidal shape (24). In such experiments it was proven that for specified solutions the melting temperature depends on the DNA sequence and is maximal for full-length perfect matches. Thus, it is possible to assess the extent of specific hybridisation and cross-hybridisation by measuring melting curves over increasing hybridisation or washing stringencies. In some early applications of microarray technologies it was pointed out that such 'melting curves could provide an additional dimension to the system and allow differentiation of closely related sequences' (25). Subsequently, similar methods were used for mutation diagnostics in the β -globin gene (26), for the determination of on-chip DNA duplex thermodynamics (27,28) and for the highly parallel study of DNA interactions with low molecular weight ligands (29) and proteins (30). However, this principle has until now not been applied to the most popular application of microarrays, the expression profiling technology, using DNA-chips.

Here we use this method to examine probe specificity on a custom-made DNA glass chip in combination with different pools of target sequences isolated from a set of different mouse tissues. We present a novel approach providing precise information about the specificity of hybridisation for each probe (also called feature) of an array. The SAFE (specificity assessment from fractionation experiments) protocol is based on the washing of microarrays with increasing stringencies and the recording of the hybridisation signal intensity for each array element at each step. In case there are different fractions of target hybridised to the same probe, these will be washed off from the array at various stringencies due to different extents of double strand formation. The set of such data for each array element comprises the fractionation curve, which provides novel information that can be used to evaluate hybridisation data reliability. The iterative use of this approach improves the selection of gene-specific probes for DNA microarray experiments based on experimental data and will thereby optimise the overall performance of DNA microarrays.

MATERIALS AND METHODS

Tissue collection

Breeding of wild-type C3HeB/FeJ mice was done under specified pathogen-free (spf) conditions. Organs were collected at the age of 105 (\pm 5) days. To minimise the influence of circadian rhythm on gene expression, mice were killed between 9 am and noon by carbon dioxide asphyxiation.

Organs (kidney, testis, brain and seminal vesicles) were dissected, weighed, snap frozen and stored in liquid nitrogen until isolation of total RNA.

Embryos were dissected at E10.5 in ice-cold phosphate-buffered saline (PBS). Chorion tissue, yolk sack and amnion were removed. Dissected embryos were stored at -80°C until isolation of total RNA.

Isolation of total RNA

All reagents were purchased from Sigma-Aldrich, unless otherwise specified. Total RNA was isolated just before processing for expression profiling. For preparation of total RNA individual organs were thawed in buffer containing chaotropic salt (RLT buffer; Qiagen) and homogenised with a Polytron homogeniser. Total RNA from individual samples was obtained according to the manufacturer's protocols using either RNeasy Mini or Midi kits (Qiagen). The concentration of total RNA was measured by OD_{260/280} reading. Aliquots were run on a formaldehyde agarose gel to check for RNA integrity. The RNA was stored at -80°C in RNase-free water until fluorescent labelling.

Reverse transcription and fluorescent labelling

For labelling 40 μg total RNA from individual tissues was used for reverse transcription and indirect fluorescent labelling. This was done using either a fluorescence indirect labelling kit (Clontech) with minor modifications of the manufacturer's protocol or the aminoallyl labelling of RNA for microarrays following the TIGR protocol (http://atarrays.tigr.org/PDF_Folder/Aminoallyl.pdf). Modifications to the Clontech protocol included an extension of the reverse transcription reaction to at least 1 h and a final ethanol precipitation of labelled DNA at -80°C for 2 h.

Preparation of probe/clone set

The 20 000 (20K) cDNA mouse arrayTAG set (Lion Bioscience) was used to produce bacterial lysates by inoculating bacterial cultures with a 96-needle replicator. The bacteria were grown in 1 ml LB medium in the presence of 100 $\mu\text{g}/\text{ml}$ ampicillin at 37°C in 96 deep-well blocks sealed with airpore sheets (Qiagen) for 24 h in a shaker. For lysates 25 μl of the bacterial cultures was mixed with 75 μl water and incubated at 95°C for 10 min. After centrifugation at 4000 r.p.m. for 5 min, 5 μl of the lysate supernatant was used for PCR. Aliquots of 95 μl PCR master-mix were added and probes were amplified.

PCR and DNA microarrays

Probes were amplified using standard PCR protocols in a Tetrad thermocycler (MJ Research) with 37 cycles (30 s at 95°C , 30 s at 52°C and 1 min at 72°C) with 5' amino-tagged primers (forward, 5'-NH₂ GTT TTC CCA GTC ACG ACG TTG-3'; reverse, 5'-NH₂ TGA GCG GAT AAC AAT TTC ACA CAG-3'; MWG-Biotech) from the non-redundant and sequence-verified Lion mouse arrayTAG 20K clone set. PCR products were amplified to a minimum concentration of 75–100 $\mu\text{g}/\mu\text{l}$ in 99.9% of the clones. All 20 000 probes were quality checked by agarose gel electrophoresis. In the entire set only seven clones did not amplify and 10 clones showed multiple bands, confirming the high quality of this particular set of mouse clones.

Clones were dissolved in 3× SSC and spotted on aldehyde-coated slides (CEL Associates) using the Microgrid TAS II spotter (Biorobotics) with 48 Stealth™ SMP3 pins (Telechem). Spotted slides were rehydrated overnight in a humid chamber containing a 50% aqueous solution of glycerol. Rehydrated slides were dried again, immersed in blocking solution (0.1 M sodium borohydride in 0.75× PBS with 25% ethanol) for 5 min, boiled in water for 2 min, briefly immersed in 100% ethanol and air dried. Slides were stored in slide boxes at ambient temperature until hybridisation.

Hybridisation, washing and image analysis

DNA microarrays and glass coverslips (Erie Scientific) were pre-hybridised for 45 min at 42°C in pre-hybridisation buffer (6× SSC, 1% BSA, 0.5% SDS). After this pre-hybridisation the slides were rinsed in water and ethanol and air dried. Aliquots of 45 µl of hybridisation solution (40 µg of each type of labelled cDNA in 6× SSC, 0.5% SDS, 5× Denhardt's solution and 50% formamide) were placed on the slide and covered with a coverslip. This assembly was placed into a hybridisation chamber (Gene Machines) and immersed in a thermostatic bath at 42°C for 22–27 h. After hybridisation slides with coverslips were immersed in 40 ml of 1× SSC pre-warmed at the hybridisation temperature and vigorously shaken to detach the coverslips. Slides were rinsed in 1× SSC and 0.5× SSC at room temperature and placed in a Petri dish with 0.25× SSC. Slides were trimmed to the length of 46 mm.

A Gene Frame 19 × 60 mm microarray sealing spacer (AB Gene) was attached to another coverslip (Erie Scientific), immersed in 0.25× SSC in a Petri dish with the hybridised slide and pasted to it such that the slots at the top and bottom of the slide were not sealed (since the slide is 46 mm in length, it is 14 mm shorter than the coverslip, Fig. 1).

This assembly was placed into a microarray scanner (GenePix 4000A; Axon Instruments) and the image was scanned at two wavelengths (532 and 635 nm). Aliquots of 700 µl of 0.25× SSC were pipetted onto one of the unsealed edges of the slide while the excess solution was removed from the opposite unsealed side with filter paper. Then the slide was washed in the opposite direction with another 700 µl of the same solution. Further washes were done with increasing concentrations of formamide (in 3.5% steps) in the same 0.25× SSC buffer. The range of formamide concentrations was from 0 to 94.5%. After each washing the slide was incubated for 5 min and scanned again.

The scanned images of hybridised microarrays were processed with the GenePix Pro 3 image analysis software. The mean pixel intensities for each single feature obtained after each washing step were plotted versus the stringency as fractionation curves.

Quantitative real-time PCR

Differential expression of selected genes was verified by quantitative PCR (qPCR). qPCR was done using a Light Cycler (Roche) and the FastStart SYBR Green kit (Roche). In brief, 1 µg of total RNA was mixed with 1 µl 0.1 mM random nonamers in a volume of 11 µl, heat denatured for 5 min at 70°C and chilled in ice water. Aliquots of 4 µl of 5× first strand buffer (Life Technologies), 2 µl DTT (Life Technologies), 1 µl RNase inhibitor (40 U/µl; Roche), 1 µl 4dNTP mix (10 mM; Amersham Biosciences) and 1 µl

SuperScriptII (Life Technologies) were added and incubated at 42°C for at least 1 h. After the reaction, the enzyme was heat inactivated for 15 min at 70°C and the obtained cDNA diluted 1:5 with water. qPCR reactions were done by mixing 2.4 µl 25 mM MgCl₂, 2 µl primer mix (5 mM each) and 2 µl SYBR Green/enzyme mix to a total volume of 18 µl with water, transferring the solution to a microcapillary (Roche) and adding 2 µl of the cDNA template. Primers were designed to be 20 bp in length with a GC content of 55% to amplify a PCR product of a maximum of 200 bp spanning an intron whenever possible. Primers from the mouse HPRT 'housekeeping' gene were used as internal control. Cycling conditions were 10 min at 95°C for activation of the hot start *Taq* polymerase followed by 45 cycles of 20 s at 95°C, 20 s at 55°C and 10 s at 72°C each.

Sequencing and calculation of melting temperatures

Twenty-two clones/probes were selected for sequencing to enable calculation of melting temperatures. Clones were PCR amplified in the same manner as for microarray spotting and sequenced (MWG-Biotech) in both directions using the same primers.

For the calculation of melting temperatures vector sequences were excluded from the clone sequence and differential melting curves were calculated according to Poland's algorithm (31) in the implementation described by Steger (32) using the online program available at <http://www.biophys.uni-duesseldorf.de/local/POLAND/poland.html> with thermodynamic parameters (33) for 0.75 mM NaCl and 1 µM strand concentration. The temperature of the final peak on the differential melting curve was taken as the melting temperature of the clone.

RESULTS

Comprehensive assessment of fractionation curves

As a first step towards the identification of specific and non-specific probes on our 20K DNA-chip, we measured post-hybridisation signal intensities of every feature *in situ* after gradual increases in the washing stringencies. The result is a unique curve of hybridisation signal intensities depending on washing stringency conditions for each combination of an individual probe and a pool of target sequences isolated from a particular tissue. Signal intensities were recorded after washes with formamide in the range 0–94.5% in steps of 3.5%. We used formamide to manipulate washing stringencies instead of heating, since in our experimental set-up this allowed a precise control of washing stringencies (Fig. 1). The resulting set of such fractionation curves was examined by means of hierarchical clustering using the Cluster software available from <http://rana.lbl.gov/EisenSoftware.htm>. Prior to clustering, artefacts that were due, for example, to contamination with dust particles during washing were filtered.

In the experiment shown in Figure 2 a total of 8980 spotted probes produced a hybridisation signal that was sufficiently strong to be detected by the image analysis software. Microarray features that were not detected by the image processing software were not clustered. A selection of data for Cy5-labelled testis cDNA is presented in Figure 2. Of the probes, 48% showed a sharp transition from the hybridised to

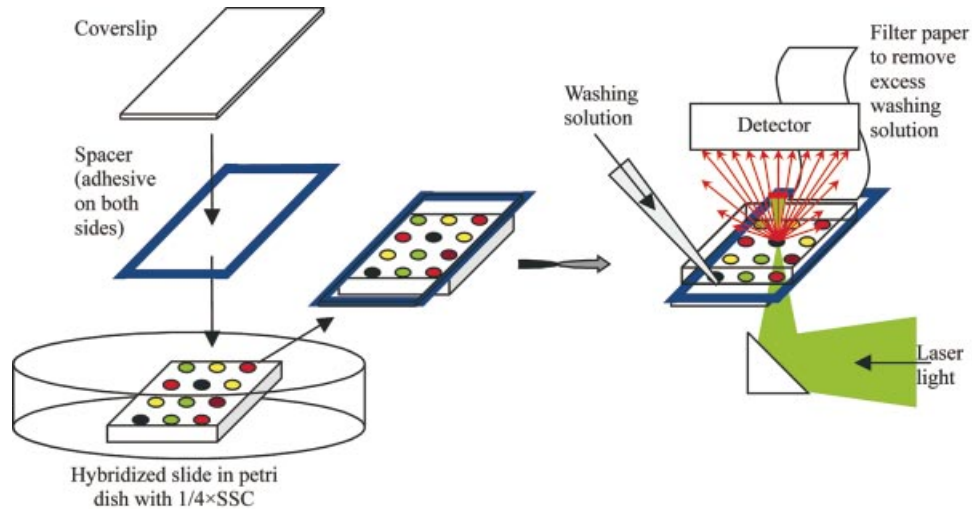


Figure 1. Scheme of experimental set-up (see Materials and Methods for description).

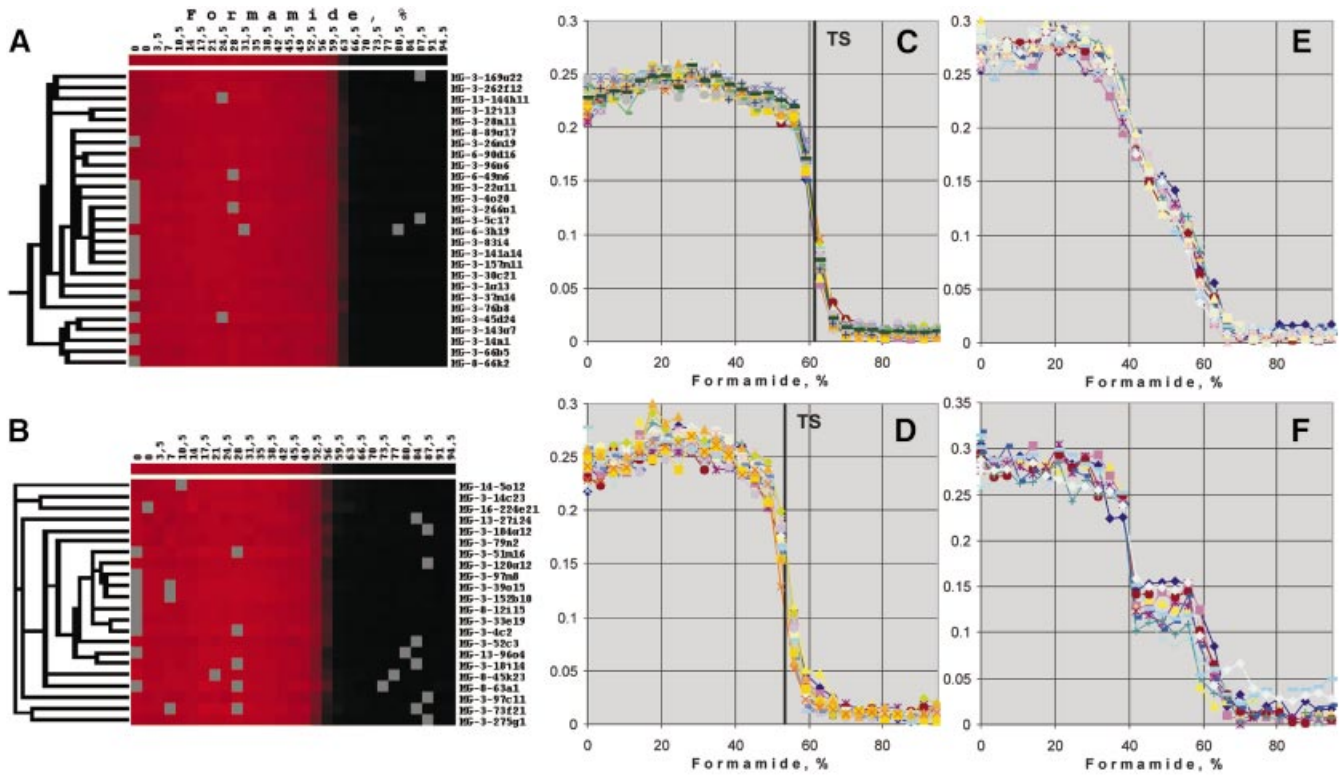


Figure 2. Comprehensive assessment of shapes of fractionation curves from normalised data. Fragments of the cluster tree representing different types of fractionation curves for Cy5-labelled testis cDNA hybridisation are shown. (A) Part of the hierarchical tree with genes having sharp transitions from the hybridised to non-hybridised state near 62% formamide that cluster together. (B) As (A) but with genes that have a sharp transition near 55% formamide. (C) Normalised signal intensities (y-axis) over increasing formamide concentrations (x-axis) of the same 27 genes as in (A). The vertical line indicates the transition stringency (TS), the mid-point of the transition from hybridised to dehybridised signal intensities. (D) Fractionation curves (x-axis, normalised signal intensities; y-axis, formamide concentration) of the same 21 genes as in (B). Vertical line indicates the transition stringency (TS) in this cluster of fractionation curves. (E) Cluster of 14 fractionation curves having broad transition regions. (F) Cluster of 10 fractionation curves having a two-step transition from the hybridised to non-hybridised state.

dehybridised state within less than 15% formamide. The stringency at which the transition occurred ranged from 40 to 70% formamide. Typical examples with transition stringencies at 62 and 55% formamide are shown in Figure 2A and C

and Figure 2B and D, respectively. For 29% of the probes the accuracy of the fractionation curves was insufficient to draw a conclusion about the character of transitions due to relatively weak signals and high noise (not shown). The remaining 23%

of clones revealed different shapes of fractionation curves, such as two-step fractionation curves (Fig. 2F), broad transition regions (Fig. 2E) and a variety of intermediate shapes (not shown).

To confirm that bleaching after repeated scans of the hybridised arrays did not significantly contribute to the fractionation curves, fluorescently labelled oligonucleotides complementary to primer sequences were hybridised to the array. After 30 scans the spot intensity was on average 72% of the initial signal intensity (not shown). Taking into account that the transition from hybridised to dissociated target molecules usually occurred over less than six scanning/washing intervals, bleaching did not significantly contribute to the shape of fractionation curves.

Based on established hybridisation behaviour in solution, we hypothesised that fractionation curves with a two-step (Fig. 2F) or broad transition (Fig. 2E) may be indicative of two or more target molecules that hybridise to these probes. In contrast, we suggest that sharp transitions (Fig. 2C and D) are a prerequisite for the hybridisation with one particular target cDNA. However, sharp transitions in signal intensities do not preclude the possibility that hybridisation is non-specific (see below).

Transition stringencies as a characteristic feature of fractionation curves

A major characteristic parameter of the fractionation curve is the transition stringency, which is defined as the mid-point of the transition region (e.g. 62% formamide for the fractionation curves in Fig. 2C, 55% formamide in Fig. 2D). Transition stringencies were highly reproducible for each probe in independent experiments, on separate DNA-chips, with different labels but from the same tissue of different individual mice. As an example, the correlation of transition stringencies (expressed as percent formamide) for kidney cDNA labelled with different fluorescent dyes and hybridised to separate slides in independent experiments is shown in Figure 3. These data have a correlation coefficient of 0.95 and a standard deviation from the best fit of 1.6% formamide. This shows that the transition stringency is a characteristic and reproducible parameter of a probe in combination with defined pools of target molecules.

Transition stringencies are more useful and efficient as criteria for evaluating probe specificity at large scale than the morphology of fractionation curves (e.g. two-step, broad and narrow transitions). On the one hand, this is due to the fact that at the moment we do not have a mathematical procedure to unambiguously classify fractionation curves based on their morphology. In addition, the precise shape of fractionation curves (in particular for two-step curves) seems to be sensitive to subtle changes in experimental conditions. For example, in a repetition of the 10 two-step curves illustrated in Figure 2F the low transition stringency step was poorly reproducible: for two probes the repetition resulted in fractionation curves that could be classified as fractionation curves with broad transition regions. Although this is compatible with the hypothesis that such a fractionation curve morphology is due to non-specific hybridisation it seems to be less reliable than an approach based on the transition stringency, which appears to be more robust (Fig. 3).

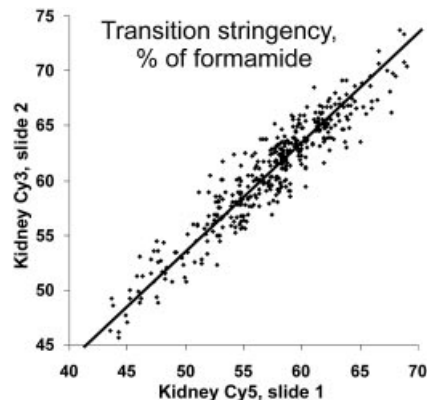


Figure 3. Transition stringencies are characteristic and reproducible parameters of a probe in combination with specific pools of target molecules. The figure shows the correlation of transition stringencies for two kidney cDNA samples, labelled with Cy3 or Cy5 and hybridised to different slides in independent experiments. The correlation coefficient is 0.95, the standard deviation from the best fit line for both Cy3 and Cy5 is 1.6% formamide. Due to the discrete values of transition stringencies in these experiments, random values with a uniform distribution from 0 to 1.5 were added to each data point, merely to avoid overlapping data points in the correlation plot. All parameters were calculated from raw data.

Transition stringencies as major criteria for probe specificity

We use the comparison of transition stringencies of individual probes in hybridisation experiments of different tissues as a measure of probe specificity. Since a full-length perfect match between probe and target is the most stable DNA duplex that can be formed, it has the maximal transition stringency. In the case of mismatched or partial hybridisation, which occurs in cross-hybridisation, the transition will take place at a lower stringency. Here we use the reduced transition stringency as an indicator of non-specific hybridisation: if for a particular clone the transition stringency is lower for the cDNA from one tissue as compared to a reference tissue and if this is confirmed in a colour flip experiment (switching the fluorescent labels), then we conclude that this clone produces non-specific hybridisation with the cDNA pool from the experimental tissue.

To compare transition stringencies and to address the question of probe specificity we hybridised a set of cDNAs isolated from different mouse tissues that is routinely used in the analysis of expression profiles from mutant mouse lines. As an example, the analysis of transition stringencies from hybridisations with cDNAs from whole embryos (E10.5) and adult testis is shown (Fig. 4). To normalise fractionation curves of individual probes we first calculated the median signal intensities for all probes on the microarray over increasing stringency (Fig. 4A and B, showing the corresponding colour flip experiments). The data shown represent the normalised median over all spots detected by the image processing software. The data were normalised by subtracting the residual signal intensities from all measuring points such that the median of the last seven measuring points (at high stringency) was set to 0. In addition, signal intensities from all measuring points were multiplied by a scaling factor such that the median signal intensities of the first seven measuring points (at low stringency) was 1. Thus, Figure 4A shows the normalised, median fractionation curve over all gene

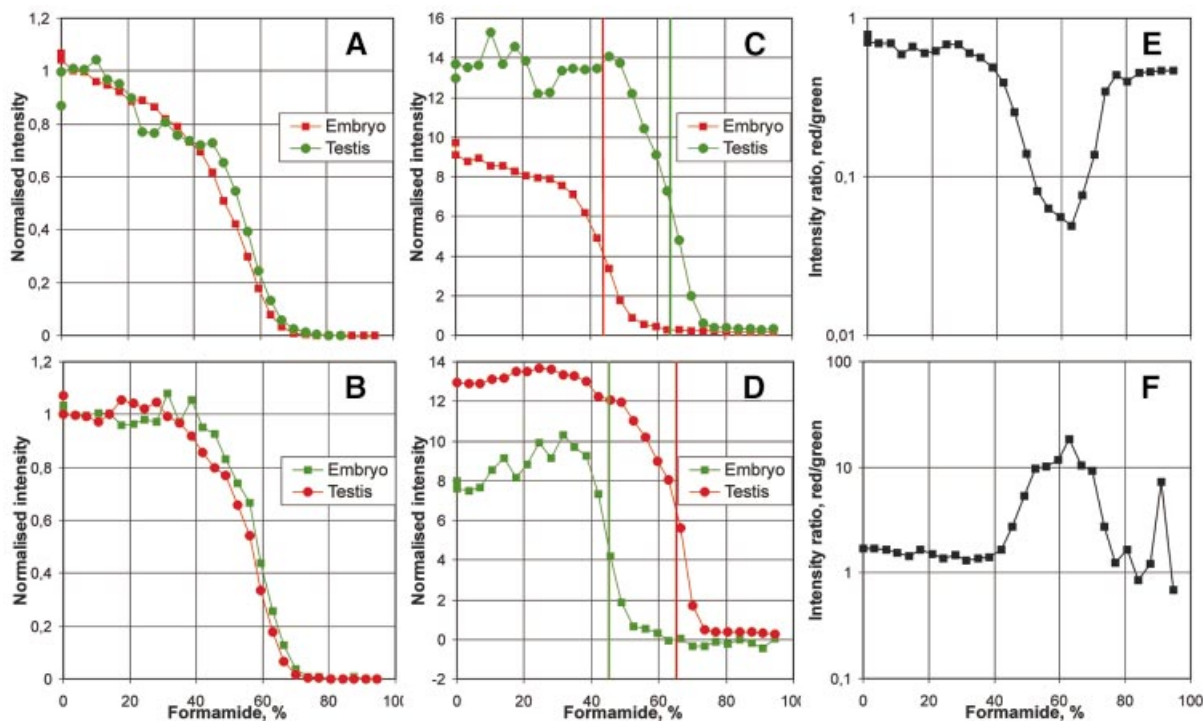


Figure 4. Using transition stringencies to determine probe specificity. Normalised fractionation curves (A–D) and ratio curves (E and F) for embryo versus adult testis hybridisation in colour flip experiments. (A) and (B) show the median of the fractionation curves for all detected spots for embryo versus testis hybridisation. The normalisation was done by subtracting the remaining signal at high stringency such that the median of the last seven measuring points was set to 0 and multiplying by a scaling factor so that the median of the first seven points at high stringency is 1. (A) Embryo, Cy5 versus testis, Cy3. (B) Embryo, Cy3 versus testis, Cy5. (C)–(F) shows the analysis of transition stringencies for one particular probe, HSP40, in the same experiments. (C) Fractionation curves of HSP40 for the hybridisation experiment shown in (A). The green curve (testis, Cy3) shows a shift of the transition region by ~20% formamide to high formamide concentrations as compared to the red curve (embryo, Cy5). The data was normalised by applying the same normalisation factors as in (A). (D) Normalised HSP40 fractionation curves for the hybridisation experiment shown in (B) (for embryo, Cy3 versus testis, Cy5). The red curve (testis, Cy5) has a shift of the transition region by ~20% formamide to high concentrations relative to the green curve (embryo, Cy3). Normalised similar to (C) with the parameters from (B). (E and F) Ratios of signal intensities measured in (C) and (D), respectively. The curves illustrate the differences in transition stringencies in the two tissues, testis and embryo, for the HSP40 gene.

expression detected in embryo (red) and testis (green). Figure 4B shows the corresponding result in the colour flip experiment. Whereas the shapes of the median fractionation curves are similar and reproducible in both tissues, we find that transition stringencies are slightly increased by approximately 2% formamide for the green fluorescent dye. A similar difference was observed in all experiments (see Fig. 3 where for Cy3-labelled target cDNAs transition stringencies were on average 3.5% higher than for Cy5). This difference is comparable to the spread of transition stringencies and is not significant for the subsequent analysis of transition stringencies of individual probes. It may be attributed to the influence of fluorescent label on DNA duplexes or, alternatively, may result from slight differences in the efficiency of signal detection for the two fluorescent dyes with our array scanner.

An example for the analysis of transition stringencies for individual probes is illustrated in Figure 4C and D for the probe corresponding to the mouse HSP40 gene. The fractionation curves for this gene were normalised by subtracting the same residual signal intensity at high stringency and multiplying by the same scaling factor as in Figure 4A and B, respectively. The data show that the HSP40 transition stringency for cDNA from embryo tissue is significantly lower (by ~20% formamide) as compared to

the transition stringency for testis cDNA (Fig. 4C). This finding was confirmed in the corresponding colour flip experiment (Fig. 4D). The initial normalised signal intensity for embryo cDNA was 60–65% of the intensity for testis cDNA in both experiments. Thus, based on the gene expression data in a normal expression profiling experiment (corresponding to the measurement at 0% formamide) it would have been estimated that HSP40 in embryo is expressed at 60–65% of the level in testis. However, the reduced transition stringency of HSP40 in embryo indicates that this signal results from extensive cross-hybridisation: at a stringency of 63% formamide the signal intensity resulting from embryo cDNA was at background level, while the decrease in the testis signal was less than half the initial signal intensity. This corresponds approximately to a 10-fold difference in the ratio of signal intensities in the transition region of the specific hybridisation in testis (63% formamide, Fig. 4E and F).

Verification of cross-hybridisation by qPCR

We used real-time qPCR to verify that expression of HSP40 in the embryo is indeed <60–65% of the expression in testis (Fig. 5). These data suggest that during the exponential phase of the PCR amplification, the background-corrected signal intensity for HSP40 in testis (Fig. 5, thick blue line) is

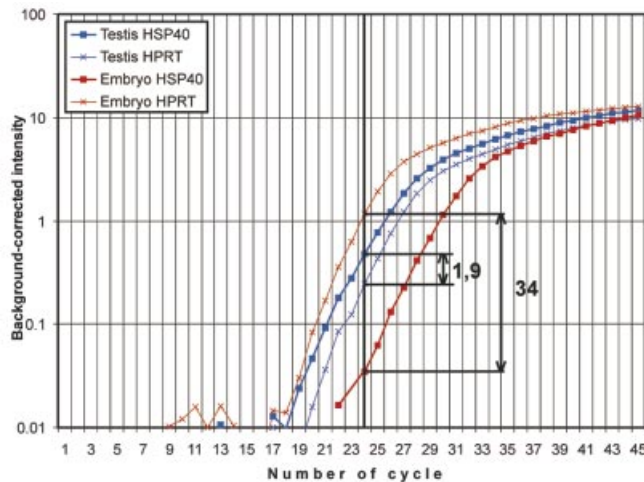


Figure 5. Quantitative real-time PCR of HSP40 and HPRT from total RNA of embryo (E10.5, brown lines) and adult testis (blue lines). The housekeeping gene HPRT was used as a reference (thin, crossed lines). In the exponential amplification phase the background-corrected (subtraction of the value corresponding to the linear signal increase at early cycles) intensity of the HSP40 gene for testis (thick blue line) was 1.9 times higher as compared to the HPRT reference (thin, crossed blue line), while for embryo it was 34 times lower (compare thick brown line and thin, crossed brown line). Thus, the differential expression of HSP40 after normalisation to HPRT is 65 times higher in testis total RNA as compared to embryo total RNA.

approximately 13 times higher than for embryo tissue (Fig. 5, thick brown line). If the data is normalised with respect to a housekeeping gene, such as HPRT (Fig. 5, thin brown and blue lines), the testis:embryo ratio for the HSP40 gene is ~65-fold. Regardless of the normalisation procedure, the real-time qPCR supports the hypothesis that expression of HSP40 in testis versus embryo is significantly higher than suggested by a standard DNA-chip experiment. Two more genes with differing transition stringencies in cDNAs from different tissues were analysed by qPCR. Based on microarray data the gene *Prkar1b* was 5-fold more strongly expressed in brain as compared to kidney. The transition stringency for the corresponding probe was significantly reduced in kidney cDNA. Accordingly, overexpression was 100-fold based on qPCR data. Similarly, the gene *Actl7b* was 2.5-fold more strongly expressed in testis as compared to embryo tissue based on microarray data and had a reduced transition stringency in fractionation curves with cDNA from embryo tissue. Overexpression was 7.1-fold based on qPCR. All qPCR results were reproduced in two or three experiments. These findings support the hypothesis that reduced transition stringencies indicate non-specific hybridisation to immobilised probes on microarrays.

Towards a comprehensive approach to estimate cross-hybridisation

To begin to comprehensively assess the specificity of the probes used on our 20K mouse DNA-chip we compared transition stringencies from total RNA isolated from a subset of organs that are routinely used in the analysis of expression profiles of mouse mutant models. The organs analysed in this

study comprise adult kidney, testis, brain, seminal vesicles and whole embryos (E10.5). To analyse fractionation curves we performed pair-wise hybridisations of these organs (Fig. 6), including the corresponding colour flip experiments. Transition stringencies were compared in both experiments, using the ratios of signal intensities over increasing stringency (as in Fig. 4E and F).

This analysis is reasonable only if the signal intensities of both fractionation curves are high at low stringencies and decrease significantly over increasing stringencies. In particular, signal intensities close to background levels would lead to division by 0 or produce high noise. Therefore, for the comparison of transition stringencies in different tissues, we selected only those probes having a mean signal intensity above a specific threshold for both wavelengths (i.e. Cy5 and Cy3). This threshold was 150 arbitrary fluorescence units for both hybridisations in experiment 1, 200 units for experiments 2 and 4 and 150 units in one hybridisation of experiment 3 and 400 units in the corresponding colour flip hybridisation of experiment 3. For example, in experiment 1 (embryo/testis) we identified 4452 genes that were expressed above this threshold in both tissues and in both corresponding colour flip experiments. 1456 such genes were identified between embryo and kidney (experiment 2), 748 between testis and seminal vesicles (experiment 3) and 3171 between brain and kidney (experiment 4) (Fig. 6, last column).

Exclusion of non-specific hybridisation

To identify probes that result from non-specific hybridisation we compared transition stringencies between tissues. As a measure for the difference in transition stringencies we evaluated the ratio curves (as in Fig. 4E and F). Each ratio curve with a peak of at least 1.4 relative to the median of the curve was verified individually. For example, in experiment 1 64 probes with a transition stringency that was significantly lower in total RNA isolated from embryo as compared to total RNA from adult testis were identified (Fig. 6, left column). In turn, for testis RNA 10 probes were identified with reduced transition stringencies as compared to embryo RNA (Fig. 6, left column). The probes listed in the left column of Figure 6 have been annotated as resulting in non-specific hybridisation in the corresponding tissue.

The limited data presented here suggests that at least 0.2% (10 of 4452, testis, experiment 1) to 1.7% (13 of 748, seminal vesicles, experiment 3) of the probes evaluated by the criteria described above produce signals that result from non-specific hybridisation. However, the portion of such non-specific probes is most likely significantly higher. It would be required to compare fractionation curves of more tissues, since transition stringencies could be decreased for both tissues used in one hybridisation experiment. As an example, in experiment 2 the transition stringency of the HSP40 gene was at 49% formamide for both embryo and kidney, while in experiment 1 it was 46% formamide for embryo and 65% formamide for testis (Fig. 4C and D). Therefore, only experiment 1 was suitable to identify the HSP40 probe as non-specific for the assessment of expression in embryo RNA.

In addition, a significant number of probes had decreased transition stringencies in one fractionation curve, while for the colour flip hybridisation the signal was too weak to determine the transition stringency (Fig. 6, middle column). This finding

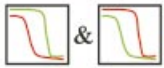
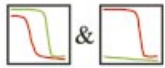

		False positives confirmed by colour flip 	False positives in one hybridisation and undetectable in another 	Spots exceeding threshold at both wavelengths in both hybridisations 
Experiment #1	Embryo Testis	64 10	354 4	4452
Experiment #2	Embryo Kidney	9 6	4 5	1456
Experiment #3	Testis Seminal vesicles	3 13	1 65	748
Experiment #4	Brain Kidney	12 12	3 6	3171

Figure 6. Summary of genes with decreased transition stringency found in different experiments. Each experiment (1–4) consists of two hybridisations (including a colour flip hybridisation) each with simultaneous hybridisation of two different tissues. The genes with decreased transition stringency (referred to as false positives) in both hybridisations are summarised in the first column for each tissue. Some genes were found to be false positives only in one experiment while in the colour flip hybridisation they produced no considerable hybridisation signal (second column). The number of features detected by the image processing software and having a mean signal across the curve above a threshold in both hybridisations is summarised in the third column for each experiment.

could be due, for example, to minor variations in hybridisation conditions. It is likely that such probes may also produce signals that result from non-specific hybridisation.

Comparison of melting temperatures and transition stringencies

It may be expected that probes with transition stringencies below a particular threshold should be considered as resulting in cross-hybridisation. To verify this, 22 probes present on our array were fully sequenced and their theoretical melting temperatures were calculated (see Supplementary Material for probe sequences). To evaluate their correlation, these melting temperatures were plotted versus their transition stringencies measured in experiment 1 (Fig. 7). Nine of the 22 selected probes had significantly different transition stringencies in testis and embryo RNA (Fig. 7, white squares, lower transition stringencies). The correlation plot from probes with equal/maximal transition stringencies in both tissues (black squares) describes a different region in the graphic (separated by dotted line) than those with reduced transition stringencies (with one exception, which is most likely due to the fact that the measured transition stringency for this probe is not maximal, similar to the low transition stringency of HSP40 in both tissues of experiment 2). However, there is a correspondence between calculated melting temperatures and the maximal measured transition stringencies (black squares, region above dotted line). This characteristic may be useful for the evaluation of the specificity of hybridisation based on the measurement of transition stringencies from single tissue RNAs and the sequence of the probe, without the measurement of transition stringencies in relation to other reference RNAs.

DISCUSSION

Although the DNA-chip technology has been applied successfully for expression profiling projects (see Introduction),

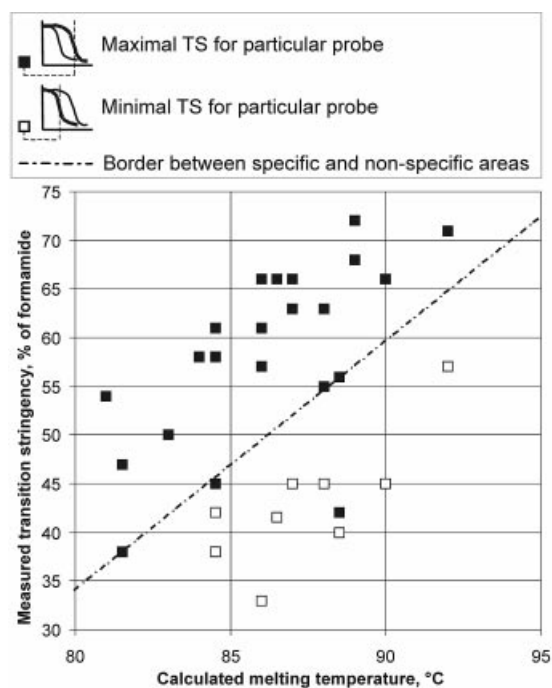


Figure 7. Correlation plot of the experimentally measured transition stringencies (testis and embryo hybridisation, experiment 1 from Fig. 6) versus the calculated melting temperatures for 22 fully sequenced probes. For nine of them the transition stringencies (TS) were different for embryo and testis RNA samples (white squares, lower TS). Other probes with the same transition stringency are indicated by black squares. The line represents the border between the areas of white and black squares, i.e. the border between non-specific and presumably specific areas.

there is an ongoing dispute concerning the quality of expression data that can be obtained from such experiments. It is known from practical experience with established hybridisation technologies, such as northern and Southern

blotting and *in situ* hybridisation methods, that the quality of the data obtained in these approaches critically depends on the selection of probes that specifically hybridise to the target mRNA. Whereas in single gene approaches it is possible to assess probe specificity empirically, this has until now not been feasible for genome wide sets of probes. Theoretical considerations such as avoiding repetitive sequences and conserved functional domains of paralogous genes have been suggested as criteria for the selection of specific probes. The applicability of this strategy depends on the completeness of sequence information. Another approach, used also for the clone set in the study described here, utilises probes that are preferentially derived from 3' untranslated regions. Using the SAFE protocol, we provide here, for the first time, a method to assess probe specificity on a large scale based on experimental hybridisation data.

Technically, expression profiling using DNA-chips is similar to the procedures of the classical dot-blot: gene-specific oligonucleotides or double-stranded cDNAs are immobilised as probes in defined positions on a solid support and hybridised to complex mixtures of expressed nucleic acids. Using the current standards for microarray spotters, up to 50 000 spots may be fitted on a standard chip of the size of a common histological slide. An important advantage of using glass as a transparent, solid support is that it allows the simultaneous, competitive hybridisation of test and reference samples labelled with different fluorescent dyes. Relative expression levels are analysed directly by comparing each fluorescent signal on every feature. An additional advantage of the DNA-chip technology, as compared to other expression profiling methods such as SAGE (serial analysis of gene expression), is that the production, hybridisation and scanning of such DNA-chips can be automated to a great extent, allowing for high throughput approaches.

The hybridisation specificity of probes depends on the population of target molecules that compete for hybridisation with the nucleotide sequence of the probe and on the stringency condition that is used in the experiment. A probe that produces a specific signal in a hybridisation experiment with total RNA from one tissue may show extensive cross-hybridisation with total RNA from another tissue that expresses other populations of genes. We demonstrate that reduced transition stringencies determined in fractionation curves of simultaneous hybridisation experiments with RNAs from different tissues are indicative of non-specific hybridisation signals.

This tissue-related information about the probe specificity is an efficient tool to validate data on differentially expressed candidate genes based on attributed weights or confidence in the probe. Using the experimental set-up described here, the measurement of fractionation curves on DNA-spotted glass slides takes ~5 h for a single hybridisation experiment. To fully implement the validation of probe specificities based on fractionation curve data it would be required to measure transition stringencies in a combinatorial way using a considerable set of different RNA pools. For example, we apply the DNA-chip technology to systematically analyse expression profiles of a selection of 17 mouse organs in a compendium of several hundred established mouse mutant lines (34). The comprehensive assessment of transition stringencies in this set of RNA pools would require the

experimental measurement of 136 pairs of tissues in at least two experiments (i.e. the corresponding colour flip hybridisations). The further automation of measuring fractionation curves and developing algorithms to analyse transition stringencies would make it feasible to estimate probe specificities on DNA-chips on a large scale.

Such comprehensive analyses of fractionation curves will result in the identification of reliable probes for expression profiling studies using the DNA-chip technology. This approach could ultimately be used to identify reliable probes for each gene that result in high quality expression data in a wide range of RNA pools from different resources. The data presented here (in particular in Fig. 6) provide a first step towards this goal. To complete this data set we are currently developing reliable software tools for the calculation of transition stringencies from fractionation data. Although the data presented were produced exclusively with immobilised PCR amplified probes, the described procedures should in principle be as useful for oligonucleotide arrays. The application to oligonucleotide arrays may, however, require the adjustment of experimental parameters such as hybridisation conditions.

In addition, we provide evidence that transition stringencies that result from specific hybridisation signals (maximal transition stringencies) correlate well with the calculated melting temperature of the corresponding probe sequence (Fig. 7). Thus, the comparison of the experimentally measured transition stringency with the calculated melting temperature of a full-length hybridisation with the probe provides an additional means to estimate potential probe specificity. In contrast to the full experimental approach described above, this method does not rely on measuring differences between diverse RNA pools. Instead, the transition stringency measured in a single experiment may be compared to the theoretical melting temperature to assess probe specificity.

The correlation of melting temperatures and formamide stringencies at which the transition from hybridised to non-hybridised target molecules occurs is a phenomenological observation that we made in the course of this study. Although such a correlation may have been expected (35), an adequate physical model does not underline it. It implies that an increase in temperature during washing steps has the same effect as an increase in stringency by elevating formamide concentrations. It also does not take into account that melting temperatures are calculated for dsDNA in solution, whereas fractionation curves are measured with probes that are immobilised on a solid surface. Although the influence of these factors may not be significant for measuring transition stringencies in the majority of cases, a proper physical model should be elaborated. Alternatively, the accuracy of fractionation curve measurements could be further improved by detecting signal intensities *in situ* during washing conditions with increasing temperature instead of formamide concentrations. However, this is not possible with currently available microarray scanners and would require considerable changes in the technological set-up.

The SAFE protocol described here provides a novel tool for the assessment of probe specificity used in genome wide DNA-chip expression profiling experiments. These procedures will allow the selection of specific probes that will lead to

high quality expression profiling data resulting from DNA-chip experiments.

SUPPLEMENTARY MATERIAL

Supplementary Material is available at NAR Online.

ACKNOWLEDGEMENTS

We would like to thank Michael Schulz and Sandra Schädler for technical assistance and Oliver Thulke for helping to solve persistent technical problems with the microarray spotter. We are grateful to Jerzy Adamski for critical reading of the manuscript. This work was supported by a grant from the German Human Genome Project (DHGP) and BFAM, Bioinformatics for the Analysis of Mammalian Genomes.

REFERENCES

- Lipshutz,R.J., Fodor,S.P., Gingeras,T.R. and Lockhart,D.J. (1999) High density synthetic oligonucleotide arrays. *Nature Genet.*, **21**, 20–24.
- 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. *et al.* (1996) Expression monitoring by hybridization to high-density oligonucleotide arrays. *Nat. Biotechnol.*, **14**, 1675–1680.
- Brown,P.O. and Botstein,D. (1999) Exploring the new world of the genome with DNA microarrays. *Nature Genet.*, **21**, 33–37.
- 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.
- Beckers,J. and Hrabe de Angelis,M. (2002) Large-scale mutational analysis for the annotation of the mouse genome. *Curr. Opin. Chem. Biol.*, **6**, 17–23.
- Hughes,T.R., Marton,M.J., Jones,A.R., Roberts,C.J., Stoughton,R., Armour,C.D., Bennett,H.A., Coffey,E., Dai,H., He,Y.D. *et al.* (2000) Functional discovery via a compendium of expression profiles. *Cell*, **102**, 109–126.
- White,K.P., Rifkin,S.A., Hurban,P. and Hogness,D.S. (1999) Microarray analysis of *Drosophila* development during metamorphosis. *Science*, **286**, 2179–2184.
- Elek,J., Pinzon,W., Park,K.H. and Narayanan,R. (2000) Relevant genomics of neurotensin receptor in cancer. *Anticancer Res.*, **20**, 53–58.
- Dhanasekaran,S.M., Barrette,T.R., Ghosh,D., Shah,R., Varambally,S., Kurachi,K., Pienta,K.J., Rubin,M.A. and Chinnaiyan,A.M. (2001) Delineation of prognostic biomarkers in prostate cancer. *Nature*, **412**, 822–826.
- Pomeroy,S.L., Tamayo,P., Gaasenbeek,M., Sturla,L.M., Angelo,M., McLaughlin,M.E., Kim,J.Y., Goumnerova,L.C., Black,P.M., Lau,C. *et al.* (2002) Prediction of central nervous system embryonal tumour outcome based on gene expression. *Nature*, **415**, 436–442.
- Heller,R.A., Schena,M., Chai,A., Shalon,D., Bedilion,T., Gilmore,J., Woolley,D.E. and Davis,R.W. (1997) Discovery and analysis of inflammatory disease-related genes using cDNA microarrays. *Proc. Natl Acad. Sci. USA*, **94**, 2150–2155.
- Beckers,J., Hoheisel,J., Mewes,W., Vingron,M. and Hrabe de Angelis,M. (2002) Molecular phenotyping of mouse mutant resources by RNA expression profiling. *Curr. Genomics*, **3**, 121–129.
- Nadeau,J.H., Balling,R., Barsh,G., Beier,D., Brown,S.D., Bucan,M., Camper,S., Carlson,G., Copeland,N., Eppig,J. *et al.* (2001) Sequence interpretation. Functional annotation of mouse genome sequences. *Science*, **291**, 1251–1255.
- Lee,C.K., Weindruch,R. and Prolla,T.A. (2000) Gene-expression profile of the ageing brain in mice. *Nature Genet.*, **25**, 294–297.
- Lee,C.K., Klopp,R.G., Weindruch,R. and Prolla,T.A. (1999) Gene expression profile of aging and its retardation by caloric restriction. *Science*, **285**, 1390–1393.
- Sandberg,R., Yasuda,R., Pankratz,D.G., Carter,T.A., Del Rio,J.A., Wodicka,L., Mayford,M., Lockhart,D.J. and Barlow,C. (2000) Regional and strain-specific gene expression mapping in the adult mouse brain. *Proc. Natl Acad. Sci. USA*, **97**, 11038–11043.
- Porter,J.D., Khanna,S., Kaminski,H.J., Rao,J.S., Merriam,A.P., Richmonds,C.R., Leahy,P., Li,J. and Andrade,F.H. (2001) Extraocular muscle is defined by a fundamentally distinct gene expression profile. *Proc. Natl Acad. Sci. USA*, **98**, 12062–12067.
- Livesey,F.J., Furukawa,T., Steffen,M.A., Church,G.M. and Cepko,C.L. (2000) Microarray analysis of the transcriptional network controlled by the photoreceptor homeobox gene *Crx*. *Curr. Biol.*, **10**, 301–310.
- Campbell,W.G., Gordon,S.E., Carlson,C.J., Pattison,J.S., Hamilton,M.T. and Booth,F.W. (2001) Differential global gene expression in red and white skeletal muscle. *Am. J. Physiol.*, **280C**, 763–768.
- Knight,J. (2001) When the chips are down. *Nature*, **410**, 860–861.
- Lee,M.L., Kuo,F.C., Whitmore,G.A. and Sklar,J. (2000) Importance of replication in microarray gene expression studies: statistical methods and evidence from repetitive cDNA hybridizations. *Proc. Natl Acad. Sci. USA*, **97**, 9834–9839.
- Tseng,G.C., Oh,M.K., Rohlin,L., Liao,J.C. and Wong,W.H. (2001) Issues in cDNA microarray analysis: quality filtering, channel normalization, models of variations and assessment of gene effects. *Nucleic Acids Res.*, **29**, 2549–2557.
- Li,C. and Wong,W.H. (2001) Model-based analysis of oligonucleotide arrays: expression index computation and outlier detection. *Proc. Natl Acad. Sci. USA*, **98**, 31–36.
- Voet,D. and Voet,J.G. (1995) Nucleic acids, structures and manipulation. In Rose,N. (ed.), *Biochemistry*, 2nd Edn. John Wiley & Sons, New York, NY, pp. 862–863.
- Stimpson,D.I., Hoijer,J.V., Hsieh,W.T., Jou,C., Gordon,J., Theriault,T., Gamble,R. and Baldeschwieler,J.D. (1995) Real-time detection of DNA hybridization and melting on oligonucleotide arrays by using optical wave guides. *Proc. Natl Acad. Sci. USA*, **92**, 6379–6383.
- Drobyshev,A., Mologina,N., Shik,V., Pobedimskaya,D., Yershov,G. and Mirzabekov,A. (1997) Sequence analysis by hybridization with oligonucleotide microchip: identification of beta-thalassemia mutations. *Gene*, **188**, 45–52.
- Kunitsyn,A., Kochetkova,S., Timofeev,E. and Florentiev,V. (1996) Partial thermodynamic parameters for prediction stability and washing behavior of DNA duplexes immobilized on gel matrix. *J. Biomol. Struct. Dyn.*, **14**, 239–244.
- Fotin,A.V., Drobyshev,A.L., Proudnikov,D.Y., Perov,A.N. and Mirzabekov,A.D. (1998) Parallel thermodynamic analysis of duplexes on oligodeoxyribonucleotide microchips. *Nucleic Acids Res.*, **26**, 1515–1521.
- Drobyshev,A.L., Zasedatelev,A.S., Yershov,G.M. and Mirzabekov,A.D. (1999) Massive parallel analysis of DNA-Hoechst 33258 binding specificity with a generic oligodeoxyribonucleotide microchip. *Nucleic Acids Res.*, **27**, 4100–4105.
- Krylov,A.S., Zasedateleva,O.A., Prokopenko,D.V., Rouviere-Yaniv,J. and Mirzabekov,A.D. (2001) Massive parallel analysis of the binding specificity of histone-like protein HU to single- and double-stranded DNA with generic oligodeoxyribonucleotide microchips. *Nucleic Acids Res.*, **29**, 2654–2660.
- Poland,D. (1974) Recursion relation generation of probability profiles for specific-sequence macromolecules with long-range correlations. *Biopolymers*, **13**, 1859–1871.
- Steger,G. (1994) Thermal denaturation of double-stranded nucleic acids: prediction of temperatures critical for gradient gel electrophoresis and polymerase chain reaction. *Nucleic Acids Res.*, **22**, 2760–2768.
- Blake,R.D. and Delcourt,S.G. (1998) Thermal stability of DNA. *Nucleic Acids Res.*, **26**, 3323–3332.
- Hrabe de Angelis,M.H., Flaswinkel,H., Fuchs,H., Rathkolb,B., Soewarto,D., Marschall,S., Heffner,S., Pargent,W., Wuensch,K., Jung,M. *et al.* (2000) Genome-wide, large-scale production of mutant mice by ENU mutagenesis. *Nature Genet.*, **25**, 444–447.
- Blake,R.D. and Delcourt,S.G. (1996) Thermodynamic effects of formamide on DNA stability. *Nucleic Acids Res.*, **24**, 2095–2103.