A dual-tag microarray platform for high-performance nucleic acid and protein analyses

Olle Ericsson[1], Jonas Jarvius[1], Edith Schallmeiner[1], Mathias Howell[1], Rachel Yuan Nong[1], Hendrik Reuter[2], Meinhard Hahn[2], Johan Stenberg[1], Mats Nilsson[1] and Ulf Landegren[1,*]

[1]Department of Genetics and Pathology, The Rudbeck Laboratory, Uppsala University, SE-75185 Uppsala, Sweden and [2]Division of Molecular Genetics, B060, Deutsches Krebsforschungszentrum, Im Neuenheimer Feld 280, D-69120 Heidelberg, Germany

ABSTRACT

DNA microarrays serve to monitor a wide range of molecular events, but emerging applications like measurements of weakly expressed genes or of proteins and their interaction patterns will require enhanced performance to improve specificity of detection and dynamic range. To further extend the utility of DNA microarray-based approaches we present a high-performance tag microarray procedure that enables probe-based analysis of as little as 100 target cDNA molecules, and with a linear dynamic range close to 10^5. Furthermore, the protocol radically decreases the risk of cross-hybridization on microarrays compared to current approaches, and it also allows for quantification by single-molecule analysis and real-time on-chip monitoring of rolling-circle amplification. We provide proof of concept for microarray-based measurement of both mRNA molecules and of proteins, converted to tag DNA sequences by padlock and proximity probe ligation, respectively.

INTRODUCTION

The DNA microarray technology introduced in the early 1990s allows rapid and cost-effective acquisition of genetic information (1,2). It has evolved towards greater densities of arrayed probes, and the probes used have largely shifted from long cDNA clones (3) to synthetic oligonucleotides. More recently, assay procedures have been developed to translate large sets of target sequences into tag sequence elements that can be sorted on general-purpose tag microarrays (4–6). The approach has been used for analysis of single-nucleotide variation (7), mRNA expression (8), and to investigate splice patterns (5).

Recently, doubts about the reliability of hybridization-based microarray analyses (9,10) have motivated nucleic acid analyses by massively parallel sequencing as an alternative to microarray hybridization (11). Massively parallel sequencing provides improved fidelity by calling targets via sequencing, and precise quantitative measurements are achieved by counting individual molecules. However, throughput and cost of parallel sequencing is not yet competitive with those of microarrays. Real-time PCR also improves specificity of detection over that of hybridization microarrays, and molecules can be measured over very wide concentration ranges, but only single or few sequences can be investigated in each reaction, limiting throughput (12). A comparison of nucleic acid analysis techniques is outlined in Supplementary Data Table 2.

The padlock probes used in the present study, and in particular a variant thereof called molecular inversion probes, have been applied for parallel analyses of up to tens of thousand target nucleic acid sequences in a single reaction, with excellent sequence specificity due to the two target-complementary regions of each individual probe that become joined by ligation in the presence of specific target sequences (13). In a similar manner, proteins can be measured with so-called proximity probes, yielding DNA reporter strands that accurately and sensitively reflect the presence of target proteins (14). We have now developed a dual-tag microarray (DTM) platform, in order to ensure that the specificity of detection and dynamic range of padlock and proximity probing reactions is not lost due to cross reactions when amplification products of reacted probes are sorted on tag arrays.

*To whom correspondence should be addressed. Tel: +46 18 471 4910; Fax: +46 18 471 4808; Email: ulf.landegren@genpat.uu.se

The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors
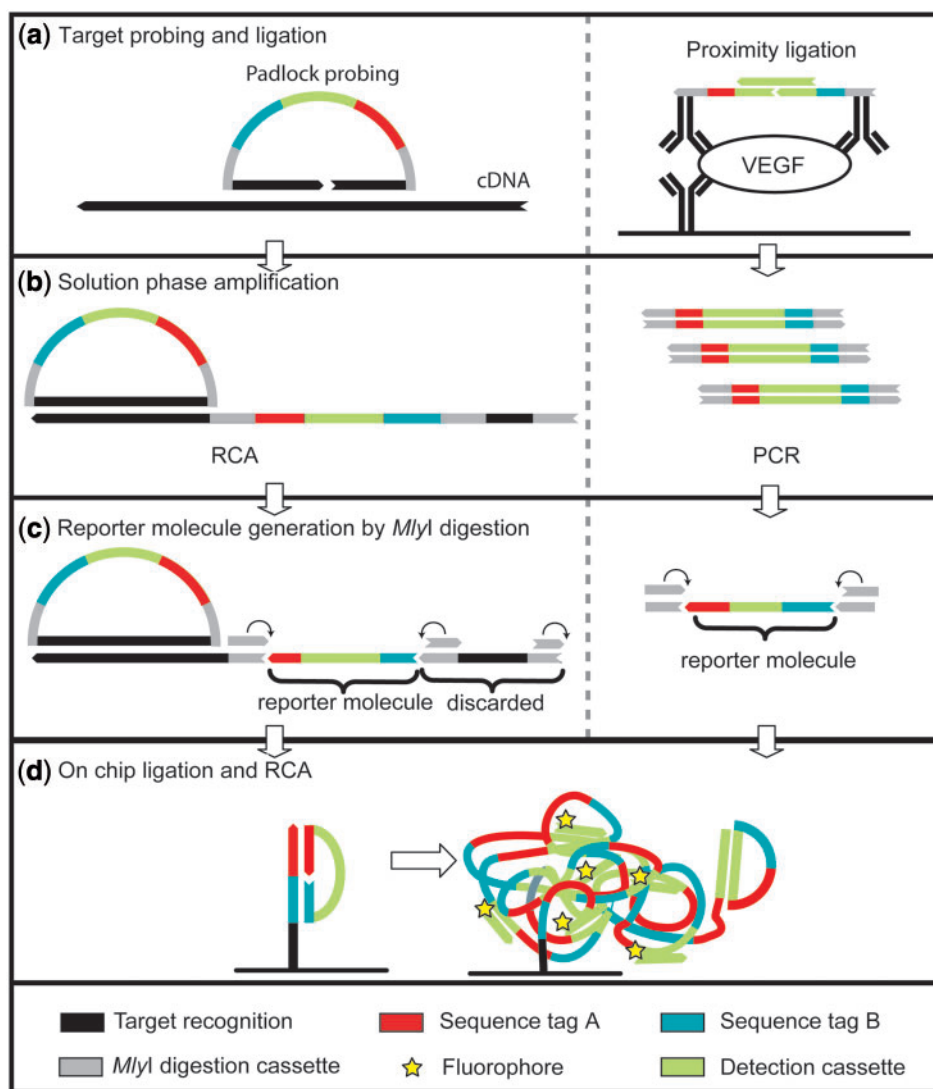
Figure 1. Schematic description of the dual-tag array procedure. (**a**) Probe design. Padlock probes are designed to comprise end sequences complementary to adjacent target sequences (black), MlyI restriction digestion cassettes (gray), and 3′ and 5′ tag sequences flanking a general detection cassette (blue, red and green, respectively). For proximity ligation each antibody is equipped with an oligonucleotide comprising a MlyI cassette, a tag sequence and half the detection sequence, which is also used for probe ligation. (**b**) The probes are ligated by target-mediated ligation. Solution phase amplification is performed and (**c**) restriction oligonucleotides are annealed to the amplification products directing the MlyI digestion, which creates reporter molecules with specific tag sequence elements at the 5′ and 3′-ends. (**d**) The reporter molecules are circularized by on-chip ligation after hybridization to microarray oligonucleotides, complementary to the tag sequences at either end of the reporter molecules. The immobilized oligonucleotides then prime on-chip RCA amplification detected by hybridizing a fluorophore-labeled detection oligonucleotide.

In the DTM procedure target nucleic acid or protein molecules are first recognized in solution by probes that include pairs of tag DNA sequences (Figure 1a). Next, ligated probes are amplified (Figure 1b) and digested in solution to form single-stranded DNA reporter molecules with one tag sequence at each end (Figure 1c). Upon hybridization of the amplified molecules to arrayed oligonucleotides, the reporter molecules are then circularized by a process of on-chip ligation, followed by on-chip rolling circle amplification (RCA) and detection (Figure 1d). RCA (15) has previously been demonstrated to enable selective amplification of reacted padlock probes (16) and the amplification level may be extended by introducing consecutive rounds of RCA (17). In a landmark

paper Lizardi *et al.* demonstrated solid-phase RCA and enumeration of single RCA products (18).

The method enhanced the limit of detection by a factor of ~100 000 compared to direct target hybridization and permitted measurement of less than 100 target molecules with a linear dynamic range close to five orders of magnitude in analyses of nucleic acid RNA expression. We observed a dramatic reduction of cross-hybridization on arrays compared to hybridization-based microarray analysis, enabling detection over wide concentration ranges. The performance of the DTM approach was additionally illustrated by digitally counting individual RCA products in array features for high-performance analysis, by monitoring on-chip RCA signal accumulation

in real-time, and by using the arrays for recording the detection of a cytokine protein via proximity ligation.

MATERIALS AND METHODS

cDNA synthesis

Total RNA was prepared from the cell lines KAT-4 (19) and MP6 (20) using the PARIS kit (Ambion Inc., Austin, TX, USA). The mRNA was converted to cDNA using oligo-dT primer and the Omniscript kit (Qiagen Nordic, West Sussex, UK) according to the manufacturer's instructions.

Padlock probe detection with DTM

The DTM procedure for padlock probe analyses was conducted using the following consecutive steps:

(1) Padlock ligation reactions were performed in $10 \mu l$ reactions, comprising $0.1 nM$ of each probe, $5 \mu l$ sample and $5 U$ of Ampligase in Ampligase buffer (Epicentre Biotechnologies, WI, USA) for four cycles of $4 h$ at $50°C$, separated by $2 min$ at $95°C$.

(2) After detection probe ligation, $1 \mu l$ of $100 nM$ RCAprim1 (Supplementary Data Table 2) was added and RCA was performed at $37°C$ for $1 h$ by addition of $5 \mu l$ S8 buffer [$20 mM$ TRIS–Ac pH 8, $50 mM$ KAc, $10 mM$ $(NH_4)_2SO_4$, $10 mM$ $MgAc_2$, $1 mM$ DTT], with $0.1 \mu g/\mu l$ bovine serum albumin (BSA) (New England Biolabs, USA), $0.6 mM$ of each dNTP and $3 U$ phi29 DNA polymerase (Fermentas, Lithuania).

(3) RCA products were cleaved with MlyI by addition of $5 \mu l$ S8 buffer with $0.1 \mu g/\mu l$ BSA, $10 pmol$ of each of the two restriction oligonucleotides 3′MC and 5′MC (Supplementary Data Table 2), along with $5 U$ of MlyI (New England Biolabs, USA) for $1 h$ at $37°C$.

(4) On-chip ligation was performed by adding $30 \mu l$ ligation mix, comprising Ampligase, buffer (Epicentre Biotechnologies, WI, USA), $25 \mu g$ BSA and $5 U$ Ampligase and incubating at $50°C$ over night after addition to the microarray reaction chambers (see Supplementary Data for details). Subsequent to ligation on arrays, each well was gently flushed with $1 ml$ of $0.75×$ TNT buffer at RT ($110 mM$ NaCl, $7.5 mM$ Tris–HCl pH 8, 0.04% Tween 20), before the microarrays were washed and dried.

(5) Next, the circularized reporter molecules were replicated by on-chip RCA in $50 \mu l$ phi29 buffer (New England Biolabs, MA, USA), $0.5 \mu g/\mu l$ BSA, $0.2 mM$ of each dNTP and $10 U$ of phi29 polymerase for $45 min$, and then the microarrays were washed and dried.

(6) Amplification products were visualized by hybridizing $10 nM$ of detection probe labeled with Cy5 (Supplementary Data Table 2) in $2×$ SSC ($0.3 M$ NaCl, $30 mM$ Na-citrate) and 0.1% SDS for at least $1 h$ at $50°C$, followed by washing and drying.

Unless otherwise stated, all microarray washes were performed in $0.75×$ TNT buffer for $5 min$, rinsing in $0.1×$ SSC, and the slides were dried by centrifugation in a table centrifuge. Microarrays were scanned with a Genepix scanner (Molecular Devices Corporation, CA, USA) at $5 \mu m$ resolution and analyzed with the Genepix Pro 6.0 software. Data points are mean values of triplicate array features, unless otherwise noted.

Protein detection with DTM

Tubes (AJ Roboscreen GmbH, Germany) were coated with $1 ng/\mu l$ affinity purified polyclonal antibody directed against vascular endothelial growth factor (VEGFA) (R&D systems Cat. # AF-293-NA) overnight at $4°C$ and the oligonucleotides ProxA and ProxB (Supplementary Data Table 2) were conjugated to two aliquots of polyclonal antibody as previously described (21). Coating of tubes with the same polyclonal antibody was followed by blocking [2% BSA (w/v), $10 ng/\mu l$ PolyA and $1 mM$ biotin]. After sample incubation for $3 h$ at $37°C$ in PBS-T [$1×$ PBS 0.05% (v/v) Tween 20], $1 nM$ of each of the two antibody oligonucleotide conjugates were added in PBS-T and incubated for $1 h$ at $37°C$. All steps were separated by washing in PBS-T in a Columbus plate washer (Tecan). Finally, a PCR and ligation mix was added [$1×$ PCR buffer (Invitrogen), $3 mM$ $MgCl_2$, $0.08 mM$ ATP, $100 nM$ ProxSplint oligonucleotide, $100 nM$ of each primer ProxFwd and ProxRev (Supplementary Data Table 2), $1 U$ T4 DNA ligase and $1.5 U$ *Taq* DNA polymerase Platinum (Invitrogen)]. The ProxFwd primer had a $5'$ phosphate, while ProvRev was blocked with a $5'$ biotin. After PCR ($2 min$ $95°C$ and 26 cycles of $15 s$ $95°C$, $1 min$ $60°C$) the buffer was changed to lambda exonuclease buffer [$67 mM$ glycine–KOH pH 9.4, $2.5 mM$ $MgCl_2$, 0.01% (v/v) Triton X-100] using a G-50 gel-filtration column (GE Healthcare). Next, ssDNA was generated by addition of lambda exonuclease, selectively destroying the strand extended from the phosphorylated but not from the biotinylated amplification primer. After heat deactivation at $80°C$ for $15 min$, and a second G-50 exchange to S8 buffer the protocol was performed according to step 3 and forward in the DTM protocol described earlier with the exception of using the oligonucleotides ProxDet and ProxTag (Supplementary Data Table 2) for detection and array tag, respectively.

RESULTS

The DTM procedure for measuring nucleic acids and proteins is illustrated in Figure 1. Briefly, for mRNA expression profiling cDNA served as targets for enzymatic circularization of padlock probes that were then amplified by RCA in solution. Similarly, in protein analysis templates for array-based RCA were obtained through proximity ligation reactions, as described later. The amplification products of probes having detected DNA or protein molecules were digested to release reporter molecules with specific tag motifs at both the 5′- and 3′-ends. These amplified reporter molecules exclusively contained *in silico* designed sequence motifs, since the digestion removed any target-specific sequences. Upon hybridization of both ends of the reporter molecules to complementary arrayed oligonucleotides, the DNA
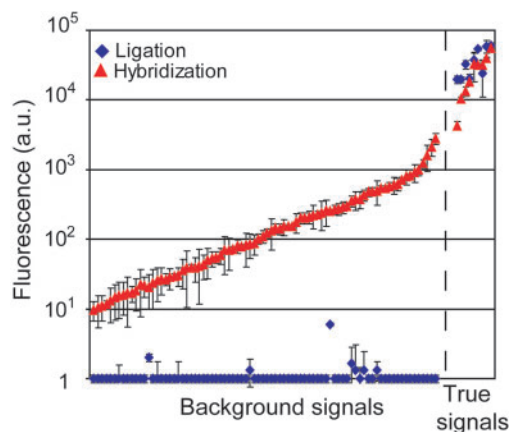
Figure 2. Paired tag ligation and hybridization signals from nine true signals and sequence-matched background signals from a high-density oligonucleotide microarray. Features are sorted according to ascending hybridization signals on the *x*-axis. Fluorescence is plotted in arbitrary units (a.u.). In total, 82 background features (to the left of the dashed line) along with the 9 true signal features (to the right of the dashed line) are plotted.

Figure 3. Evaluation of DTM performance. Measurement of signal amplification in DTM-based detection. Comparison of: sandwich hybridization of a pre-circularized reporter molecule and a detection oligonucleotide (squares), on-chip RCA of a pre-circularized reporter molecule and hybridization of a detection probe (triangles) and padlock probe-based detection of a synthetic template and amplification by RCA in solution followed by on-chip RCA and visualization with a detection probe (diamonds). The concentration of the nucleic acid target is plotted on the *X*-axis and the corresponding fluorescence output on the *Y*-axis.

fragments were circularized by on-chip ligation and amplified by on-chip RCA. Finally, the on-chip RCA products in arrays were detected by hybridization using universal fluorescence-labeled oligonucleotides. The dependence on specific ligation of probes followed by amplification, first in solution and then again on-chip, maintains the chain of specificity from the detection of target molecules in solution to the recording of detection signals on the arrays.

It is well-established that DNA detection by ligation of pairs of probe sequences can provide increased specificity of detection compared to simple hybridization reactions (22). To investigate the effect on microarray performance by on-chip ligation versus hybridization, nine circulariz-able reporter molecules were used. A set of 2507 arrayed oligonucleotide probes that could be expected to exhibit cross-reactivity for the reporter molecules were identified from a commercial set of 30 000 70-mer oligonucleotide array probes by a low-stringency BLAST search (for details see Supplementary Data). This oligonucleotide set was used to evaluate microarray signal to noise in experiments where the reporter molecules were either hybridized as pre-circularized reporter molecules or circularized by on-chip ligation, templated by the immo-bilized oligonucleotides, followed by on-chip RCA. The RCA products were detected using the same labeled detection oligonucleotides in all experiments. Among the 2507 array probes, a total of 82 reproducible background signals were recovered above the noise threshold. All 82 cross-reactions appeared in the hybridization experiment while in the ligation experiment all features were below the background threshold, as seen in Figure 2 (see Supplementary Data for details). It should be pointed out that the on-chip RCA facilitates detection of cross-hybridization that might go undetected in conventional hybridization arrays because of weaker detection signals.

In order for the reporter molecules from DNA or protein detection reactions to be circularized on arrays
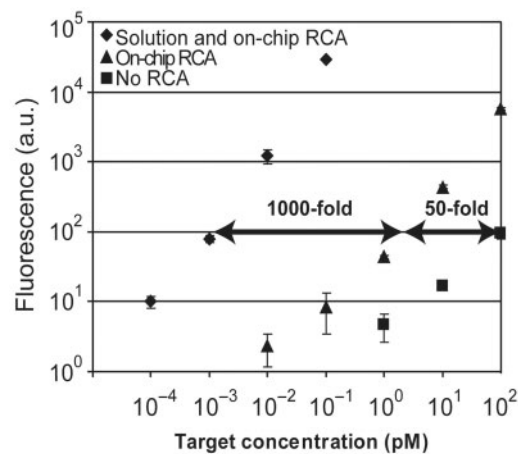
they must be single-stranded and have tag sequences at both ends. We found that the type IIS restriction enzyme MlyI was able to cleave single-stranded amplification products of ligated probes next to their variable tag sequences, as long as the immediately adjacent MlyI recognition sequences were made double-stranded (Figure 1c and Supplementary Data Figure 1). The efficiency of the MlyI cleavage at the junctions of single- and double-stranded DNA was 81% of that obtained by RsaI digestion of the double-stranded molecule (Supplementary Data Figure 1), as investigated by separating radiolabeled molecules by PAGE gel electro-phoresis (data not shown).

To benchmark the complete protocol we first investi-gated how the limit of detection was affected by the two sequential RCA steps, comparing the relative intensity of signals when; (i) pre-circularized reporter molecules hybridized to array features were directly detected by sandwich hybridization of fluorescence-labeled oligonu-cleotides without any amplification (Figure 3, no RCA), (ii) pre-circularized reporter molecules were detected by on-chip RCA (Figure 3, on-chip RCA) and (iii) using the complete procedure where circularized padlock probes were first amplified by RCA in solution, followed by restriction digestion, on-chip ligation and a second round of RCA (Figure 3, solution and on-chip RCA). The results show that around 100-fold lower concentrations of cir-cularized padlock probes could be detected using on-chip RCA, compared to direct detection of the pre-circularized reporter molecules without amplification. By combining a first solution-phase RCA with a secondary on-chip RCA, approximately 10^5-fold lower concentrations of target strands could be detected compared to the detection of the circular DNA strands by sandwich hybridization with a
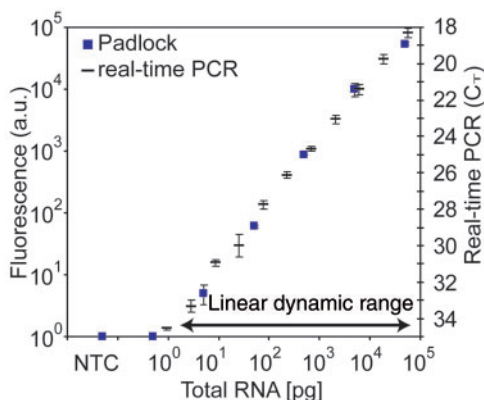
Figure 4. Comparison of beta actin cDNA detection using quantitative-PCR and padlock probes with dual-tag array. Cycle threshold (C_T) values from the quantitative-PCR are shown on the right Y-axis, and fluorescence intensity readings from the array scanner are indicated on the left Y-axis.

Figure 5. Ratio profiling of dilutions of beta actin transcripts. The indicated ratios of transcripts were investigated using padlock probes with distinct reporter sequence tags in separate reactions.

Figure 6. Two replicate analyses interrogating ten loci in four transcripts are plotted against each other to demonstrate reproducibility.

detection oligonucleotide—a situation similar to that of standard hybridization microarrays.

For a practical demonstration of the DTM procedure we compared the detection of beta actin transcripts in dilutions of total cellular cDNA, using either padlock probes with DTM or traditional quantitative real-time PCR analysis of the same cDNA dilutions (Figure 4). With the DTM approach the beta actin transcript could be detected over a linear dynamic range close to 10^5 and a limit of detection of 1.6 pg total mRNA corresponding to as little as ~10 aM or ~30 copies in 5 µl of the beta actin transcript. The limit of detection, two standard deviations above background, of the DTM procedure was calculated by scanning the microarray at higher photomultiplier tube gain to determine the standard deviation of the background signal (see Supplementary Data Figure 2). The combination of padlock probes and DTM thus allows parallel mRNA measurements with performance characteristics approaching those achieved in individual quantitative real-time PCR analyses, where even single nucleic acid copies can be detected and dynamic ranges of up to 10^7–10^8 are possible.

The ability of the DTM platform to resolve small copy number differences was investigated in a dual-color experiment, in a manner similar to comparisons between gene expression in healthy and diseased tissue samples. One set of samples contained a constant amount of total cDNA (A), and these were compared to a second set of samples (B) with a series of 1.8-fold dilutions of the same cDNA sample. For this purpose two padlock probes with identical target-complementary sequences but different detection tags were used for the A and B samples, respectively. One padlock probe, whose amplification products on the arrays were detected with probes labeled with the fluorophore TAMRA, was used for target detection in the samples with a fixed concentration (A). The other padlock probe, reporting in Cy5, was applied for target detection in the dilution series (B), in a series of separate ligation reactions. The A and B ligation reactions were subsequently pooled pair-wise, forming a single
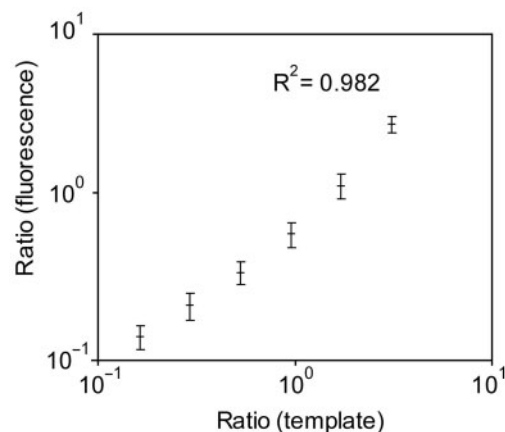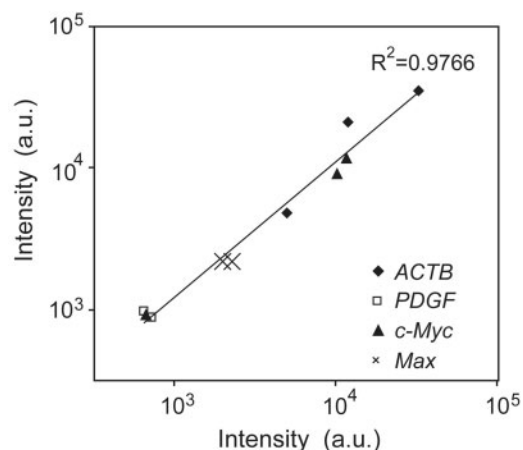
series where the ratios of the two types of reacted padlock probes with separate detection backbones reflect the target ratios in the pooled pairs of samples. The reactions were subsequently detected with DTM as outlined in Figure 1, using one TAMRA and one Cy5 labeled detection oligonucleotide for detection of RCA products, thereby reporting the target ratio in each microarray feature (see Supplementary Data Figure 3 for an illustration of the experimental procedure). The measured ratios of fluorescence from samples with greater than 1.5-fold differences in copy numbers differed by more than two standard deviations, permitting precise copy number measurement (Figure 5).

To investigate the generality and reproducibility of the DTM read-out, 10 padlock probes were designed to detect four transcripts, and their expression levels were analyzed in cDNA prepared from the KAT4 cell line. The correlation of two independent experiments showed a linear regression of 0.9766, demonstrating good reproducibility (Figure 6). Furthermore, the KAT4 cell line,
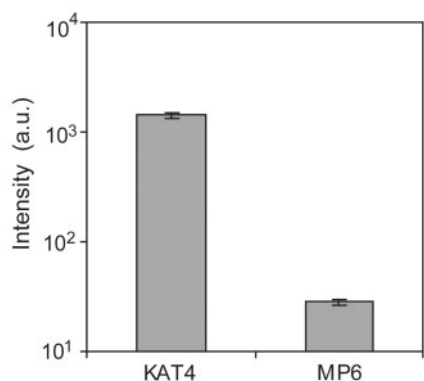
Figure 7. Comparison of PDGF transcript expression in the KAT4 and MP6 cell lines.
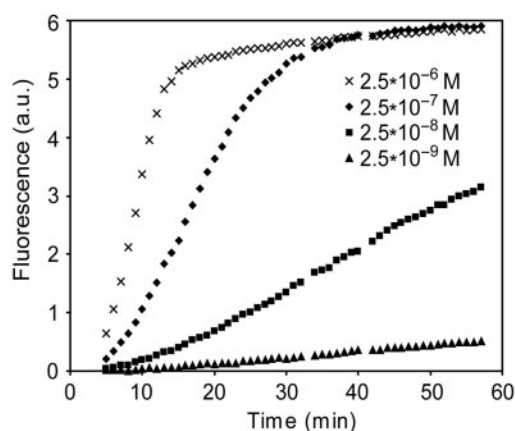

Figure 8. Real-time monitoring of accumulating fluorescence at array features during array primed RCA. Results are shown for array features printed with four different probe concentrations.

known to express high levels of transcripts for platelet-derived growth factor A (PDGFA) (19), was compared to the MP6 cell line by analyzing the summed signals of the two PDGFA probes, revealing that the MP6 cell line expressed 50-fold lower levels of PDGFA transcripts compared to the KAT4 cells (Figure 7).

We next monitored the on-chip RCA in real-time using a confocal microarray scanner. A molecular beacon probe (23), modified for detection of RCA products (24) (see Supplementary Data Table 2), was used to follow the accumulation of RCA products on microarrays during replication. For this experiment, variable amounts of oligonucleotide probes were deposited on a microarray, and an excess of pre-circularized reporter molecules was allowed to hybridize to the immobilized primers, and washed before initiating the on-chip RCA. Figure 8 demonstrates that the on-chip RCA from oligonucleotides printed on the array at four different concentrations resulted in a linear accumulation of fluorescence for up to 60 min, and that higher primer concentrations resulted in proportionately higher fluorescence from the array features. A movie illustrating the progress of signal increase in this experiment can be found in Supplementary Data.

Single-molecule detection can provide improved quantitative precision. To investigate the suitability of the DTM platform for single-molecule analysis of reporter molecules on arrays we enumerated individual amplification products within individual array features as a measure of two different oligonucleotide sequences detected by separate padlock probes. The probe pair, designed to discriminate a single-nucleotide variation (G/T), was added to samples containing two synthetic target sequences at different proportions, followed by probe ligation. The ligated padlock probes were then hybridized to array features containing an oligonucleotide complementary to a common sequence element. On-chip RCA was initiated from the array, and the two classes of amplification products were detected by hybridization of two detection oligonucleotides labeled with distinct fluorophores. The Cy3-labeled detection probe D3 reported the G-variant of the target sequence, while the Cy5-labeled detection probe D5 reported the T-variant. Images from individual microarray spots were recorded using confocal microscopy, followed by image analysis and automated counting of fluorescent spots using a Matlab script. A thresholded two-color image of a microarray feature is shown in Figure 9a. The measured ratios of discrete single-molecule counts from the two probes observed in individual array features were plotted against the ratios of the two added sequence variants, demonstrating excellent correlation. The linear regression value of the plots was 0.9935, and the average deviation from the mean ratio of the data points was 2% (Figure 9b).

Finally, the DTM platform could also be useful for recording the detection of proteins in multiplex. To investigate the feasibility of this we applied the DTM to monitor the outcome of solid-phase proximity ligation assays for protein detection (Figure 1) (25). Two oligonucleotides, one having a free 3′-end and the other a free 5′-end, were conjugated to separate aliquots of a polyclonal antibody directed against vascular endothelial growth factor A (VEGFA). Each oligonucleotide included a tag sequence and a PCR primer site, also encompassing the MlyI cleavage cassette. Following capture of the VEGFA target protein in microtiter wells by an immobilized antibody and washing, the pair of conjugates was added. Binding by appropriate pairs of antibody-oligonucleotide conjugates to individual VEGFA molecules permitted ligation of the attached oligonucleotides that were thus brought in proximity. The probe ligation products, comprising both tag sequences, were then selectively amplified by PCR. One of the PCR primers was phosphorylated, permitting degradation of the corresponding strand by lambda exonuclease (26), while the other strand was protected by a 5′ biotin moiety. The single-stranded product was cleaved by addition of restriction oligonucleotides and MlyI, leaving a reporter molecule with tags at both the 3′- and 5′-ends. The reporter molecules were circularized by on-chip ligation after hybridization to the tag microarray, amplified by RCA, and visualized by detection probe decoration. As seen in Figure 10, this on-chip analysis of products of proximity ligation reactions enabled detection of low picomolar concentrations of recombinant VEGFA protein in a format suitable for
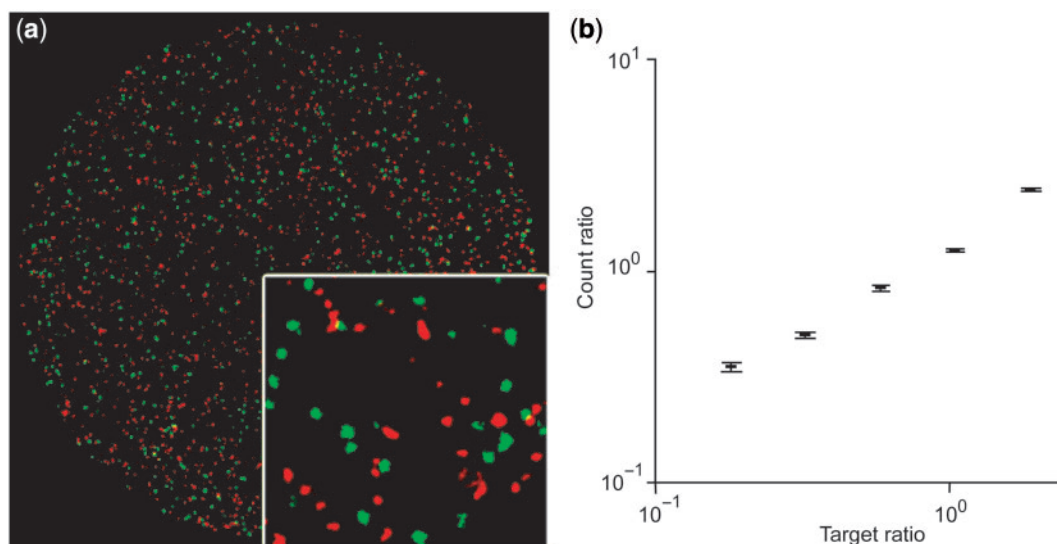
Figure 9. Single molecule counting in microarray features. (**a**) The image illustrates a microarray feature with amplification products detected using two different fluorophores after thresholding of fluorescence. Individual RCA products are featured in a magnified subregion. (**b**) Measurement of two single-nucleotide DNA sequence variants by digital detection of amplification products in individual microarray features. The ratios of the two added probe ligation templates correlated well with the single molecule count ratios from microarray features.
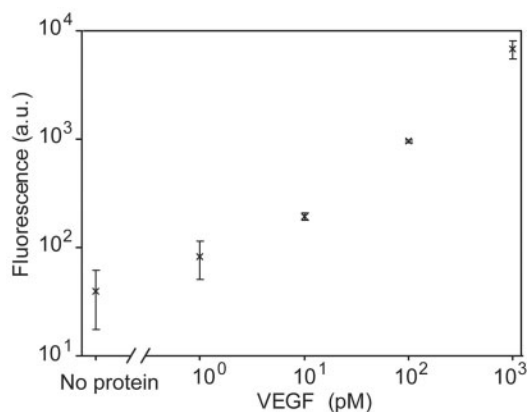

Figure 10. Application of DTMs for sensitive protein measurement of VEGFA by proximity ligation.

multiplex protein detection with high sensitivity and over wide dynamic ranges.

DISCUSSION

Microarrays provide efficient measurement of gene expression, as confirmed by recent reports where microarray platforms were validated using real-time PCR (27,28). However, there is need for improvements of e.g. false negative rates, and the dynamic range in microarray analyses is typically limited to ~3 logs (27,29,30) even though RNA expression in homogenous cell populations span over six orders of magnitude (12). Cross-hybridization that amount to less than 1–10% of the hybridization signal from the intended microarray feature is often considered insignificant (13,31–33). Nonetheless, this level may well contribute to frequent reports of poor correlation between weak microarray signals and real-time PCR measurements of transcript levels (30,34).

Padlock probes are capable of distinguishing even single base-pair differences in total human genomic DNA, and their ability to be combined in large numbers in individual reactions has proven useful for parallel SNP genotyping, measurements of gene expression, and for detection of sets of pathogens (6,35,36). The specificity introduced by ligation is well-documented (22). However, after target detection by ligation, amplification products of reacted probes are typically sorted on microarrays by hybridization, a step where specificity may be lost due to cross-hybridization to incorrect array features. Herein we present a protocol in which also microarray sorting of amplified detection probes was achieved via ligation, and we demonstrate that this significantly reduced the risk of cross-hybridization, and greatly improved detection sensitivity.

The DTM platform presented herein will allow highly specific measurements of large sets of biomolecules with a very low limit of detection, wide dynamic range and high precision. Non-specific background is minimized by using highly specific probing steps, and by ensuring that detection signals are exclusively dependent on probe ligation reactions. A first solution-phase amplification of ligated detection probes serves to enhance the kinetics of the subsequent array hybridization, and amplified reporter molecules are sorted with excellent specificity via paired sequence recognition and on-chip ligation, followed by a second amplification step on the array. Using the DTM for recording beta actin gene expression, we successfully detected less than 100 transcripts, and with a dynamic range of almost five orders of magnitude, in a protocol suitable for parallel detection of very large sets of amplified probes. By allowing the use of arrayed oligonucleotides to template ligation of reporter molecules, the limit of detection by on-chip RCA could be enhanced

without a concomitant increase of non-specific background. The improved specificity obtained by on-chip ligation was demonstrated by comparing the background signals generated by nine reporter molecules in array-based ligation or hybridization. The requirement for ligation eliminated all background signals recorded among the 2507 investigated array probes present in the corresponding hybridization-based experiment. The single- and dual-color expression profiling experiments demonstrated that the DTM approach is applicable for measuring target molecules with high precision and reproducibility. Analogously, for protein measurement by proximity ligation, reacted probes were amplified by PCR prior to array-based on-chip ligation and RCA, enabling detection of VEGFA in the low picomolar range, and in a manner suitable for multiplex protein analyses.

The opportunity to perform single-molecule quantification in individual array features can confer advantages over conventional measurements of ensemble averages. By counting discrete reporter molecules as individual, easily detectable RCA products, analytic variation can be brought to the statistical minimum defined by the Poisson distribution of numbers of molecules (37). On-chip RCA of single, specifically reacted probes as used herein avoids problems of distinguishing weak fluorescent signals from background and it ensures highly specific target detection, in contrast to other single-molecule detection approaches that rely on simple probe-binding reactions and detection of very few fluorophores over any background arising from media or nonspecifically bound probes.

In conclusion, we have shown that the DTM procedure can be used to record reacted probes that report the amounts of specific nucleic acids or proteins, with excellent performance. The procedure should also be useful to identify interacting protein pairs among large sets of proteins recognized by antibodies with attached DNA strands, by decoding which pairs of tag sequences that have become joined in the proximity ligation products using DTMs. The DTM platform thus presents an attractive alternative for advanced microarray analysis by enhancing the limit of detection and dynamic range, eliminating cross-hybridization, and by providing a means to measure individual or interacting proteins on DNA microarrays.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

Conflict of interest statement. U.L, M.N. and J.J. are founders of Olink Biosciences, a company that is commercializing the padlock probe and proximity ligation technologies.

REFERENCES

1. Southern,E.M., Maskos,U. and Elder,J.K. (1992) Analyzing and comparing nucleic acid sequences by hybridization to arrays of oligonucleotides: evaluation using experimental models. *Genomics*, **13**, 1008–1017.
2. Fodor,S.P., Read,J.L., Pirrung,M.C., Stryer,L., Lu,A.T. and Solas,D. (1991) Light-directed, spatially addressable parallel chemical synthesis. *Science*, **251**, 767–773.
3. 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.
4. Shoemaker,D.D., Lashkari,D.A., Morris,D., Mittmann,M. and Davis,R.W. (1996) Quantitative phenotypic analysis of yeast deletion mutants using a highly parallel molecular bar-coding strategy. *Nat. Genet.*, **14**, 450–456.
5. Yeakley,J.M., Fan,J.B., Doucet,D., Luo,L., Wickham,E., Ye,Z., Chee,M.S. and Fu,X.D. (2002) Profiling alternative splicing on fiber-optic arrays. *Nat. Biotechnol.*, **20**, 353–358.
6. Baner,J., Isaksson,A., Waldenstrom,E., Jarvius,J., Landegren,U. and Nilsson,M. (2003) Parallel gene analysis with allele-specific padlock probes and tag microarrays. *Nucleic Acids Res.*, **31**, e103.
7. Hardenbol,P., Banér,J., Jain,M, Nilsson,M, Namsaraev,E.A., Karlin-Neumann,G.A., Fakhrai-Rad,H., Ronaghi,M., Willis,T.D., Landegren,U. *et al.* (2003) Multiplexed genotyping with sequence-tagged molecular inversion probes. *Nat. Biotechnol.*
8. Fan,J.-B., Yeakley,J.M., Bibikova,M., Chudin,E., Wickham,E., Chen,J., Doucet,D., Rigault,P., Zhang,B., Shen,R. *et al.* (2004) A versatile assay for high-throughput gene expression profiling on universal array matrices. *Genome Res.*, **14**, 878–885.
9. Eklund,A.C., Turner,L.R., Chen,P., Jensen,R.V., deFeo,G., Kopf-Sill,A.R. and Szallasi,Z. (2006) Replacing cRNA targets with cDNA reduces microarray cross-hybridization. *Nat. Biotechnol.*, **24**, 1071.
10. Zhang,J., Finney,R.P., Clifford,R.J., Derr,L.K. and Buetow,K.H. (2005) Detecting false expression signals in high-density oligonucleotide arrays by an in silico approach. *Genomics*, **85**, 297.
11. Mikkelsen,T.S., Ku,M., Jaffe,D.B., Issac,B., Lieberman,E., Giannoukos,G., Alvarez,P., Brockman,W., Kim,T.-K., Koche,R.P. *et al.* (2007) Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature*, **448**, 553.
12. Holland,M.J. (2002) Transcript abundance in yeast varies over six orders of magnitude. *J. Biol. Chem.*, **277**, 14363–14366.
13. Hardenbol,P., Yu,F., Belmont,J., MacKenzie,J., Bruckner,C., Brundage,T., Boudreau,A., Chow,S., Eberle,J., Erbilgin,A. *et al.* (2005) Highly multiplexed molecular inversion probe genotyping: Over 10,000 targeted SNPs genotyped in a single tube assay. *Genome Res.*, **15**, 269–275.
14. Fredriksson,S., Dixon,W., Ji,H., Koong,A.C., Mindrinos,M. and Davis,R.W. (2007) Multiplexed protein detection by proximity ligation for cancer biomarker validation. *Nat. Methods*, **4**, 327–329.
15. Fire,A. and Xu,S.Q. (1995) Rolling replication of short DNA circles. *Proc. Natl Acad. Sci. USA*, **92**, 4641–4645.
16. Baner,J., Nilsson,M., Mendel-Hartvig,M. and Landegren,U. (1998) Signal amplification of padlock probes by rolling circle replication. *Nucleic Acids Res.*, **26**, 5073–5078.

17. Dahl,F., Baner,J., Gullberg,M., Mendel-Hartvig,M., Landegren,U. and Nilsson,M. (2004) Circle-to-circle amplification for precise and sensitive DNA analysis. *Proc. Natl Acad. Sci. USA*, **101**, 4548–4553.

18. Lizardi,P.M., Huang,X., Zhu,Z., Bray-Ward,P., Thomas,D.C. and Ward,D.C. (1998) Mutation detection and single-molecule counting using isothermal rolling-circle amplification. *Nat. Genet.*, **19**, 225–232.

19. Ain,K.B., Tofiq,S. and Taylor,K.D. (1996) Antineoplastic activity of taxol against human anaplastic thyroid carcinoma cell lines in vitro and in vivo. *J. Clin. Endocrinol Metab.*, **81**, 3650–3653.

20. Rosen,A., Lundman,P., Carlsson,M., Bhavani,K., Srinivasa,B.R., Kjellstrom,G., Nilsson,K. and Holmgren,A. (1995) A CD4+ T cell line-secreted factor, growth promoting for normal and leukemic B cells, identified as thioredoxin. *Int. Immunol.*, **7**, 625–633.

21. Soderberg,O., Gullberg,M., Jarvius,M., Ridderstrale,K., Leuchowius,K.J., Jarvius,J., Wester,K., Hydbring,P., Bahram,F., Larsson,L.G. *et al.* (2006) Direct observation of individual endogenous protein complexes in situ by proximity ligation. *Nat. Methods*, **3**, 995–1000.

22. Stenberg,J., Nilsson,M. and Landegren,U. (2004) In Meyers,R.A. (ed.), *Encyclopedia of Molecular Biology: Fundamentals and Applications*. VCH Publishers, Weinheim, pp. 179–193.

23. Piatek,A.S., Tyagi,S., Pol,A.C., Telenti,A., Miller,L.P., Kramer,F.R. and Alland,D. (1998) Molecular beacon sequence analysis for detecting drug resistance in Mycobacterium tuberculosis. *Nat. Biotechnol.*, **16**, 359–363.

24. Nilsson,M., Gullberg,M., Dahl,F., Szuhai,K. and Raap,A.K. (2002) Real-time monitoring of rolling-circle amplification using a modified molecular beacon design. *Nucleic Acids Res.*, **30**, e66.

25. Fredriksson,S., Gullberg,M., Jarvius,J., Olsson,C., Pietras,K., Gustafsdottir,S.M., Ostman,A. and Landegren,U. (2002) Protein detection using proximity-dependent DNA ligation assays. *Nat. Biotechnol.*, **20**, 473–477.

26. Little,J.W. (1967) An exonuclease induced by bacteriophage lambda. II. Nature of the enzymatic reaction. *J. Biol. Chem.*, **242**, 679–686.

27. Canales,R.D., Luo,Y., Willey,J.C., Austermiller,B., Barbacioru,C.C., Boysen,C., Hunkapiller,K., Jensen,R.V., Knight,C.R., Lee,K.Y. *et al.* (2006) Evaluation of DNA microarray results with quantitative gene expression platforms. *Nat. Biotechnol.*, **24**, 1115.

28. Wang,Y., Barbacioru,C., Hyland,F., Xiao,W., Hunkapiller,K.L., Blake,J., Chan,F., Gonzalez,C., Zhang,L. and Samaha,R.R. (2006) Large scale real-time PCR validation on gene expression measurements from two commercial long-oligonucleotide microarrays. *BMC Genomics*, **7**, 59.

29. Rajagopalan,D. (2003) A comparison of statistical methods for analysis of high density oligonucleotide array data. *Bioinformatics*, **19**, 1469–1476.

30. Beckman,K.B., Lee,K.Y., Golden,T. and Melov,S. (2004) Gene expression profiling in mitochondrial disease: assessment of microarray accuracy by high-throughput Q-PCR. *Mitochondrion*, **4**, 453.

31. Xu,W., Bak,S., Decker,A., Paquette,S.M., Feyereisen,R. and Galbraith,D.W. (2001) Microarray-based analysis of gene expression in very large gene families: the cytochrome P450 gene superfamily of Arabidopsis thaliana. *Gene*, **272**, 61–74.

32. Miller,N.A., Gong,Q., Bryan,R., Ruvolo,M., Turner,L.A. and LaBrie,S.T. (2002) Cross-hybridization of closely related genes on high-density macroarrays. *Biotechniques*, **32**, 620–625.

33. Chou,C.-C., Chen,C.-H., Lee,T.-T. and Peck,K. (2004) Optimization of probe length and the number of probes per gene for optimal microarray analysis of gene expression. *Nucleic Acids Res.*, **32**, e99.

34. Czechowski,T., Bari,R.P., Stitt,M., Scheible,W.-R. and Udvardi,M.K. (2004) Real-time RT-PCR profiling of over 1400 Arabidopsis transcription factors: unprecedented sensitivity reveals novel root- and shoot-specific genes. *Plant J.*, **38**, 366–379.

35. Baner,J., Marits,P., Nilsson,M., Winqvist,O. and Landegren,U. (2005) Analysis of T-cell receptor V beta gene repertoires after immune stimulation and in malignancy by use of padlock probes and microarrays. *Clin. Chem.*, **51**, 768–775.

36. Baner,J., Gyarmati,P., Yacoub,A., Hakhverdyan,M., Stenberg,J., Ericsson,O., Nilsson,M., Landegren,U. and Belak,S. (2007) Microarray-based molecular detection of foot-and-mouth disease, vesicular stomatitis and swine vesicular disease viruses, using padlock probes. *J. Virol. Methods*, **143**, 200.

37. Jarvius,J., Melin,J., Goransson,J., Stenberg,J., Fredriksson,S., Gonzalez-Rey,C., Bertilsson,S. and Nilsson,M. (2006) Digital quantification using amplified single-molecule detection. *Nat. Methods*, **3**, 725–727.