

## Cooperativity of paired oligonucleotide probes for microarray hybridization assays

Steven R. Bates<sup>a</sup>, Donald A. Baldwin<sup>b</sup>, Alexandra Channing<sup>a</sup>,  
Lida K. Gifford<sup>a</sup>, Angela Hsu<sup>a</sup>, Ponzy Lu<sup>a,\*</sup>

<sup>a</sup> Department of Chemistry, University of Pennsylvania, Philadelphia, PA 19104, USA

<sup>b</sup> Department of Pathology and Laboratory Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA

Received 23 December 2004

Available online 15 April 2005

### Abstract

Synthetic DNA probes attached to microarrays usually range in length from 25 to 70 nucleotides. There is a compromise between short probes with lower sensitivity, which can be accurately synthesized in higher yields, and long probes with greater sensitivity but lower synthesis yields. Described here are microarrays printed with spots containing a mixture of two short probes, each designed to hybridize at noncontiguous sites in the same targeted sequence. We have shown that, for a printed microarray, mixed probe spots containing a pair of 30mers show significantly greater hybridization than spots containing a single 30mer and can approach the amount of hybridization to spots containing a 60mer or a 70mer. These spots with mixed oligonucleotide probes display cooperative hybridization signals greater than those that can be achieved by either probe alone. Both the higher synthesis yields of short probes and the greater sensitivity of long oligonucleotides can be utilized. This strategy provides new design options for microarray hybridization assays to detect RNA abundance, RNA splice variants, or sequence polymorphisms.

© 2005 Elsevier Inc. All rights reserved.

**Keywords:** Oligonucleotide hybridization; Probe design

Highly parallel assays of RNA abundance are now predominantly conducted using synthetic oligonucleotide probe collections in microarray format. Array manufacturers need methods to optimize probe sensitivity and specificity given the inherent length and composition limitations of current oligonucleotide synthesis systems. Flexible probe design strategies will be important for other microarray applications, such as improved detection of sequence polymorphisms and development of assays for RNA splice variants.

Two predominant types of microarray production methods are currently in use. Photolithographic methods [1] build probes base by base in situ using masks or

micromirrors to control the deprotection chemistry at selected spots prior to adding the desired nucleotide. Printed microarrays [2] are made by depositing machine-synthesized probes onto their intended substrate locations, where they are attached covalently or through ionic interactions. Each method has advantages and disadvantages. For probes synthesized using photolithography on the array substrate, each round of base addition has a yield of about 92–94% [3] compared with 99% [4] typical for oligonucleotides synthesized by standard automated instruments before printing. The former method limits the probe length to 25–40 bases, since as the probes increase in length, the proportion with random single-base deletions will accumulate (for example, the expected proportion of 25mers without deletions is  $0.93^{25} = 16\%$ ). Printed arrays, on the other hand, allow longer probe length ( $0.99^{180} = 16\%$ ), and PCR products

\* Corresponding author. Fax: +1 215 573 2159.  
E-mail address: [ponzy@sas.upenn.edu](mailto:ponzy@sas.upenn.edu) (P. Lu).

and oligonucleotides can be used. Long probes are expected to have greater sensitivity because the target would bind with greater affinity, and the probes would have a greater effective cross section for finding an accessible complementary site on the target to nucleate hybridization. In contrast, correctly chosen short probes, as short as 15 bases, are optimal for discriminating single base mismatches [5]. One mismatch in 15 bp has more impact on hybridization specificity than one mismatch in 70 bp. As a compromise, 60mers or 70mers (“longmers”) are currently the typical probe lengths for printed oligonucleotide arrays, in part to bracket the double helix nucleation sequence [6]. Photolithographic arrays bracket the uncertain location of the optimal nucleation sequence by their ability to produce very large numbers of unique small spots. This redundancy using multiple distinct probes for the same target sequence compensates for low signal that may be due to a few poorly performing probes or inaccessible sites on the target. Also, short probes are likely to pack in a spot at a higher surface density than long probes, which tend to be more flexible and able to adopt conformations that can interfere with the attachment of neighboring molecules onto the surface [7].

To provide additional flexibility among the various factors influencing probe length, we propose that a useful option would be to design microarrays on which each spot contains two probes of different sequence (“mixed probe spots”) instead of a single probe (“single probe spots”). Mixed probes would increase the effective probe length without increasing the actual oligonucleotide length, allowing some of the benefits of long probes to be transferred to photolithographic arrays. Mixed probes would also decrease cost in the case of printed arrays. It is cheaper to synthesize two  $n$ -mers than the same molar amount of one  $2n$ -mer. This opens up new probe design combinations for novel assays on both platforms.

## Materials and methods

### Probe preparation

DNA 30mer, 60mer, and 70mer probes were synthesized with a primary amino group attached either through a phosphate group and a six-carbon linker to the 5' carbon of the first base or through a phosphate group and a seven-carbon linker to the 3' carbon of the last base. Printing solutions consisting of 150 mM sodium phosphate, pH 8.5, 0.0005% *N*-lauroyl sarcosine, and either 25  $\mu$ M probe for single probe solutions or 12.5  $\mu$ M each probe for mixed probe solutions were prepared. Solutions were printed from a 384-well plate (Genetix X6004) on Codelink Activated Slides (Amersham Biosciences 300011) or Ultra-GAPS Slides

(Corning 40017) with a Gene Machines Omnigrad 100 printer and Telechem SMP4 pins. After printing, slides were stored overnight in a chamber with a saturated sodium chloride slurry at the bottom to maintain humidity. Codelink slides were blocked by incubating in 50 mM ethanolamine, 0.1 M Tris, pH 9, and 0.1% SDS for 15 min at 50 °C, rinsed with deionized water, and washed by incubating in 4 $\times$  SSC<sup>1</sup> and 0.1% SDS for 15 min at 60 °C. Finally, they were rinsed with deionized water, centrifuged to dryness, and stored in a desiccator until use. Ultra-GAPS slides were rehydrated in steam for approximately 5 s, until condensation formed, snap-dried by placing on a 95 °C hot plate for 5 s and immobilized by applying 250 mJ/cm<sup>2</sup> of UV energy with a Stratalinker UV cross-linker.

### Target preparation

Target was made by asymmetric PCR [8] with *Taq* DNA polymerase (Promega M1861), using as template pBluescript II KS+ plasmids with a *Bacillus subtilis* DapB, LysA, or PheB gene insert (American Type Culture Collection item numbers 87486, 87482, and 87483, respectively). The primer corresponding to the sense strand of the gene insert was present in 100 $\times$  excess to the other primer, to make the target predominantly single-stranded, with the sense strand in excess. The target was labeled with Cy3 by including Cy3-dCTP (Amersham Biosciences PA53021) at one-fifth the concentration of dCTP ([dCTP]=250  $\mu$ M). Long variant LysA target was prepared by reverse PCR using primers oriented to amplify the entire plasmid, placing the targeted hybridization sequences at either end of a 3969-bp product.

### Hybridization

Labeled target was hybridized to immobilized probes by applying 45  $\mu$ l of hybridization solution—5 $\times$  SSC, 0.1% SDS, 2.5  $\mu$ g/ml target, 0.1 mg/ml salmon sperm DNA (Stratagene 201190-81), 25% formamide—to each slide under a Lifter Slip (Erie Scientific 25  $\times$  60I-2-4789). Each slide was placed in a hybridization chamber (Corning 2551) and incubated in a water bath at 57 °C between 15 and 17 h. Slides were then rinsed with 4 $\times$  SSC and washed twice for 5 min each time in 2 $\times$  SSC, 0.1% SDS preheated to 42 °C, washed for 1 min in 0.2 $\times$  SSC, washed for 1 min in 0.1 $\times$  SSC, and then centrifuged dry. Fluorescence signal from each slide was read with an Axon Genepix 4000B scanner, and the image was analyzed with Genepix Pro software.

<sup>1</sup> Abbreviations used: SSC, standard saline citrate; RFU, relative fluorescence units; SNP, single nucleotide polymorphism.

## Data analysis

Spots that had low fluorescence intensity or poor morphology were “flagged” by the software during scanning; spots were then manually flagged if they appeared misshapen and were not already flagged. For each slide, the mean and standard deviation of background intensity were calculated from blank spots containing no probe. The mean background—which had an average of 80 relative fluorescence units (RFU), and a standard deviation of 25 RFU and ranged from 50 to 117 RFU—was subtracted from each spot’s intensity, and the spots that had a recentered intensity less than two times background standard deviation were also flagged. All flagged spots were excluded from further analysis, and recentered intensities were used in all subsequent calculations. For replicate spots within an array, the median intensity and standard deviation were estimated using a bootstrap method [9,10]: the data set was sampled with replacement  $n$  times, with  $n$  the size of the data set. This process was repeated to generate 1000 “bootstrap samples.” The mean of the median values of the bootstrap samples was taken to be the median intensity. The mixed probe spots were then compared to single probe spots in one of two ways. To determine whether mixed probe spots had higher intensities than single probe spots, the median intensity of each type of mixed-probe spot was compared to the median intensity of each of the two corresponding single-probe spots using a two-sample  $t$  test, with  $H_0: I_{\text{mixed}} \leq I_{\text{single}}$ ,  $H_a: I_{\text{mixed}} > I_{\text{single}}$ , and a cutoff of  $p \leq 0.01$  for rejecting  $H_0$ . To numerically compare mixed and single probe spots, the intensity ratios were calculated using an adaptation of the “median of ratios” method [9]. A set of all possible ratios was generated by dividing each of the intensities in the mixed probe data set by each of the intensities in the single probe data set, creating a data set of size  $m \times n$ , where  $m$  is the size of the mixed probe data set and  $n$  the size of the single probe data set. The median and standard deviation of this “ratio set” was estimated using the bootstrap method with 200 bootstrap samples. In addition to comparing mixed probe spots to each of the two corresponding single probe spots, this method was used to compare mixed probe spots to spots printed with a single longmer. This analysis was performed on a slide-by-slide basis, but the median ratios for each mixed probe pair were ultimately calculated for multiple slides, using a method mathematically equivalent to pooling the 200 ratio bootstrap samples for each slide and calculating the overall median and standard deviation of median.

## Results

Mixed probe pairs were designed with four variable parameters: orientation of probe attachment to the sub-

strate, gene targeted, region of the target containing the complementary sites, and spacing between the complementary sites. The Codelink slides used for these experiments have an aminoreactive surface. To explore all possible geometries, the probes were made with an amino group attached to either the 5′ or the 3′ end. The resulting four possible combinations for mixed probe pairs (Fig. 1) are both probes attached by their 3′ (3′–3′) or 5′ (5′–5′) ends, first probe attached by its 3′ end and the second by its 5′ end (3′–5′), and first probe attached by its 5′ end and the second by its 3′ end (5′–3′). The first probe in each pair is defined as the probe whose site of complementarity occurs earlier in the coding sequence of the target. To avoid confusion, 5′ and 3′ will be used to refer to the ends of probes; N and C will be used to refer to the ends of the target (“N end” corresponds to the 5′ end of the coding sequence and the amino terminus of the gene product).

The targets were amplified from plasmids containing fragments of the *B. subtilis* genes LysA, PheB, and DapB accompanied by an artificial 41-base poly(A) region; these genes were chosen because of their established use as microarray standards [11]. The targets were therefore composed of gene insert, poly(A) region, and vector in different proportions depending on design of the PCR primers; see Fig. 1 for a summary. For LysA and PheB, two groups of probe pairs targeting the N and C halves of both targets were selected (Figs. 1B and C); one group targeted to the N end was used for DapB (Fig. 1D). Each group consisted of six pairs of probes, one of which was the same for each pair within the group (fixed probe), while the other was varied (variable probe) to alter the spacing between the two targeted sites of complementarity.

Microarray slides were printed with each of the mixed probe pairs and with each of the single probes that were members of the pairs. The total molar amount of probe per spot was the same for all experiments. Labeled target was hybridized to the slide and the relative fluorescence intensity of each spot was measured. For each slide, the median intensities of all spots containing the same probe or probe pair were calculated. Fig. 2 is an example of the data from an array targeted to the N end of the LysA sequence.  $t$  tests were performed on these medians to determine whether the fluorescence intensity of mixed probe spots was significantly greater than the intensity of each of the corresponding single probe spots. To avoid confounding of results by fluorescence intensity variation among different slides,  $t$  tests were always calculated on a slide-by-slide basis and never using overall medians across multiple slides. Mixed and single probe spots were also compared by calculating the median ratio of mixed signal intensity to single signal intensity for each slide (Figs. 3 and 4; Supplemental Figs. 1–3 online). Mixed signal intensity was also compared to signal intensity from spots that had been printed with longmers target-

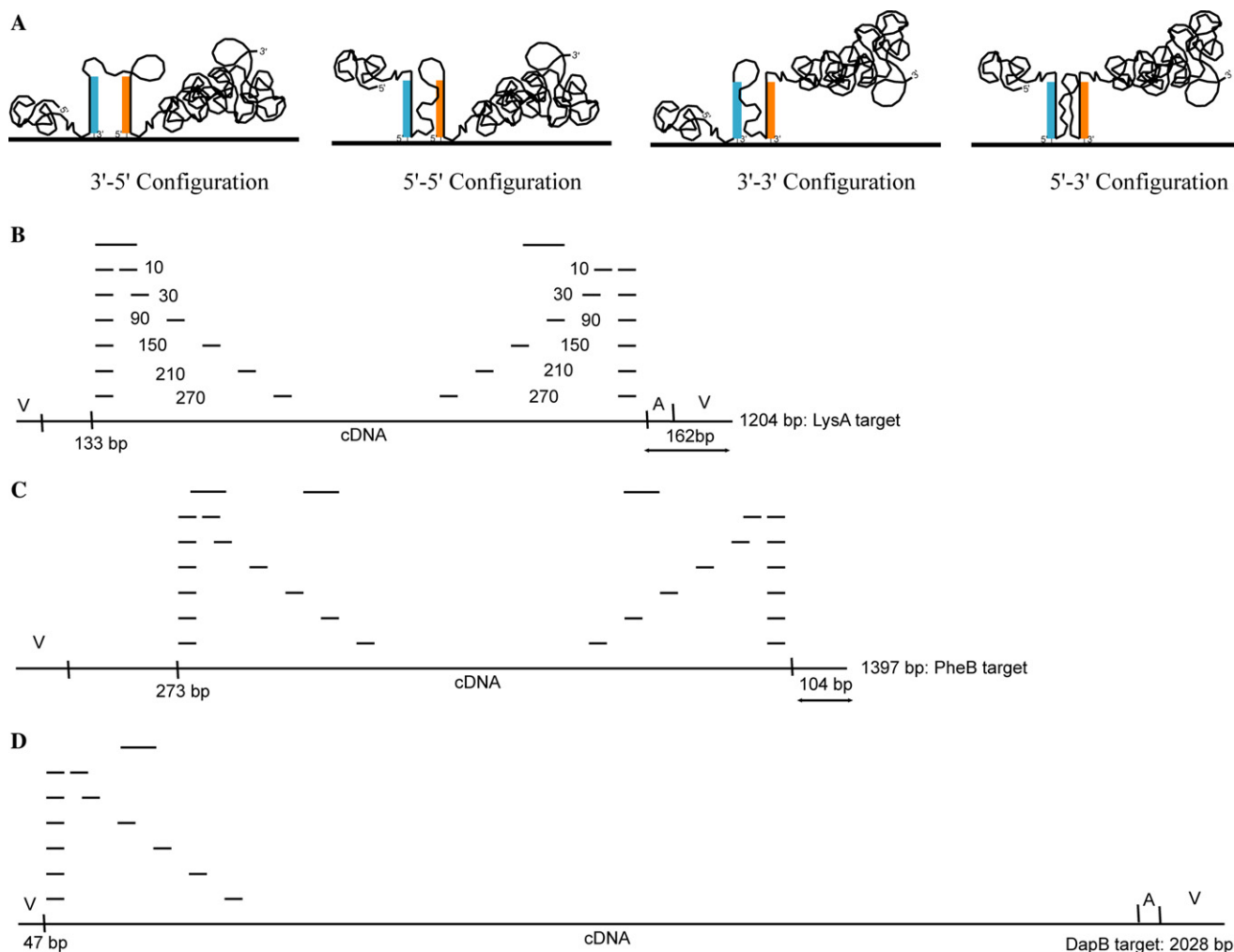


Fig. 1. (A) Cartoon, to scale of the molecular configuration at the microarray. Each of the two probes can be attached at its 3' end or 5' end, so there are a total of four possible attachment orientation combinations. These cartoons show the implications of attachment configuration on target binding. Steric interference of the target by the surface might make certain configurations more favorable. The upright shaded bars represent 30 bp or about three turns of a double helix, 10 nm in height. The random coil segments are about 5 bp or 1.7 nm. The target shown corresponds to 180 bp at the 5' end, 30 bp with the first probe, 90 bp spacing, 30 bp with the second probe, and 1095 bp on the 3' end. (B) Scale models of the three targets. Probe pairs and longmers are shown above their complementary sequences on the cDNA. V, vector sequence; A, poly(A) sequence; cDNA is the targeted gene sequence; longer dashes show the position of the 60mer probe sequences. The numbers 10, 30, 90, 150, 210, and 270 with dashes on either side are the fixed and variable probes with respective spacing (bp). The exact base sequence of each probe is in supplemental Figure 8. The probes at 133, 273, and 47 bp are the N-end fixed probes for the LysA, PheB and DapB targets respectively. The probe that ends 162 and 104 bp from the 3' ends of LysA and PheB are the C-end fixed probes, respectively.

ing the same half of the target. The signal obtained with longer probes, 60- or 70mer probes, was on average only 12% greater than that obtained with the optimum mixed pair of 30mers. This 12% difference is reproducible but does not overshadow the advantages of the mixed probes discussed here, thus *t* tests were not computed to compare mixed 30mer probe results versus 60- or 70mer probe results.

Mixed probes consistently have a higher amount of hybridization than single probes of the same length and can exhibit hybridization comparable to that of a longmer probe. As a criterion for success, mixed probes should have at least an arithmetic twofold increase in hybridization compared to corresponding single probes

and at least half the hybridization compared to longmers. Of the 120 probe pairs analyzed, 52, or 43%, fit these criteria. If the definition of success is expanded to include pairs with any increase in hybridization over single probe and more than half the hybridization compared to that of longmers, 94, or 78%, of the pairs fit the criteria.

The experiments were repeated with PheB target using Ultra-GAPS slides (Fig. 5, Supplemental Figs. 4 and 5 online), which have a surface rich in amino groups that binds probes by interacting with the phosphate backbone. Though the same aminated probes that had been printed on the Codelink slides were used, probes for the Ultra-GAPS substrate need not be func-

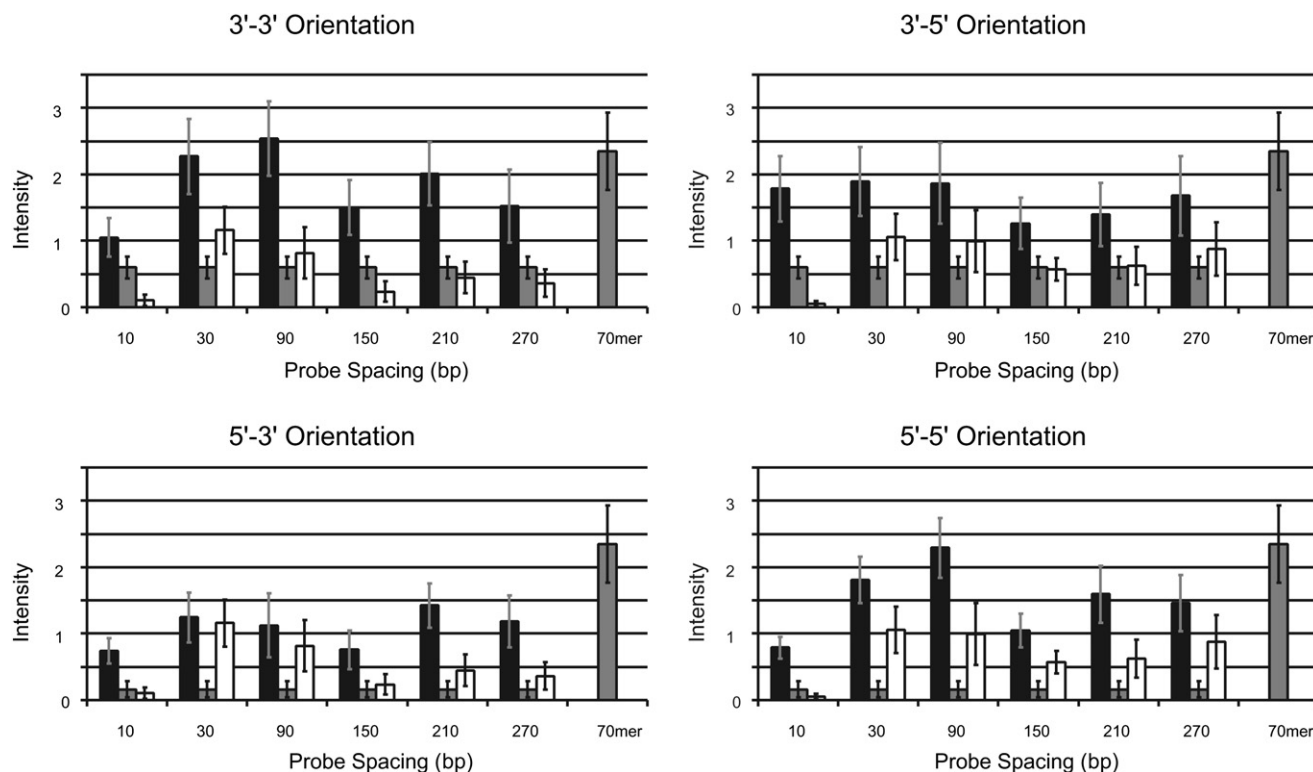


Fig. 2. Hybridization to the N end of the *LysA* target. Intensities are in relative fluorescence units (10,000 $\times$ ) and are averaged over six slides. Bars for each mixed probe pair and the two corresponding single probes are grouped. Error bars represent one standard deviation in either direction.  $x'-y'$  orientation indicates the attachment polarity of each probe. Black bar, mixed probes; dark gray bar, fixed probe; white bar, variable probe.

tionalized. Such probes should attach with a random orientation; however, repulsion between the amino groups of the probes and the substrate used in these experiments may have influenced attachment, leading to noticeable differences in mixed probe signal enhancement among the four mixed probe orientations (Supplemental Figs. 4 and 5 online). Mixed probes on Ultra-GAPS slides show a discernable signal enhancement, but not to the same degree as that on Codelink slides.

For both types of slides, single and mixed probe spots targeting the C ends did not hybridize as much or as consistently as probes targeting the N ends nor did the mixed probe spots for the C ends show appreciable signal enhancement. The amount of mixed probe enhancement was affected by probe orientation and spacing but not in a systematic way. Probe sequences and positions of the hybridization sites on the target must play a more important role in mixed probe enhancement.

To explore longer probe spacings, probe pairs targeting the *LysA* gene with spacings up to 850 bases were printed. The constant probe in these pairs was the same as the *LysA* N end constant probe. The mixed probe signal appears to drop off when the spacing is increased beyond about 500 bases (Fig. 6, Supplemental Fig. 6 online); this may be useful for estimating an unknown

distance between two targeted hybridization sites in different exons. In addition, the effect of longer target spacing on mixed spots without altering sequence complementarity was analyzed by hybridizing to the microarrays either normal *LysA* sequence or a long variant designed to separate the targeted sites by 3600–3800 nucleotides (Fig. 7). No enhancement over the signal from single probe spots was observed for the long target, suggesting that splice variants that insert sufficient length between targeted sites can be discriminated from targets with less than 500-bp spacing by a loss of cooperative signal enhancement.

Hybridization was tested for spots in which only one of the two probes in a mixed probe pair matches the target (“half-match mixed probe spots”). Probes targeting *LysA* were mixed with a probe targeting *DapB* and compared to the corresponding *LysA* single probe spots after hybridization using *LysA* target. As expected, the half-match spots had, with one exception, a decreased signal compared to single probe spots (Fig. 8), presumably because there are only half the number of probes available to bind target, although the total number of nucleic acids attached in each spot is constant. This signal ratio decrease below 1 is a confirmation of the cooperation effect of mixed probes where both complimentary sequences exist in the target.

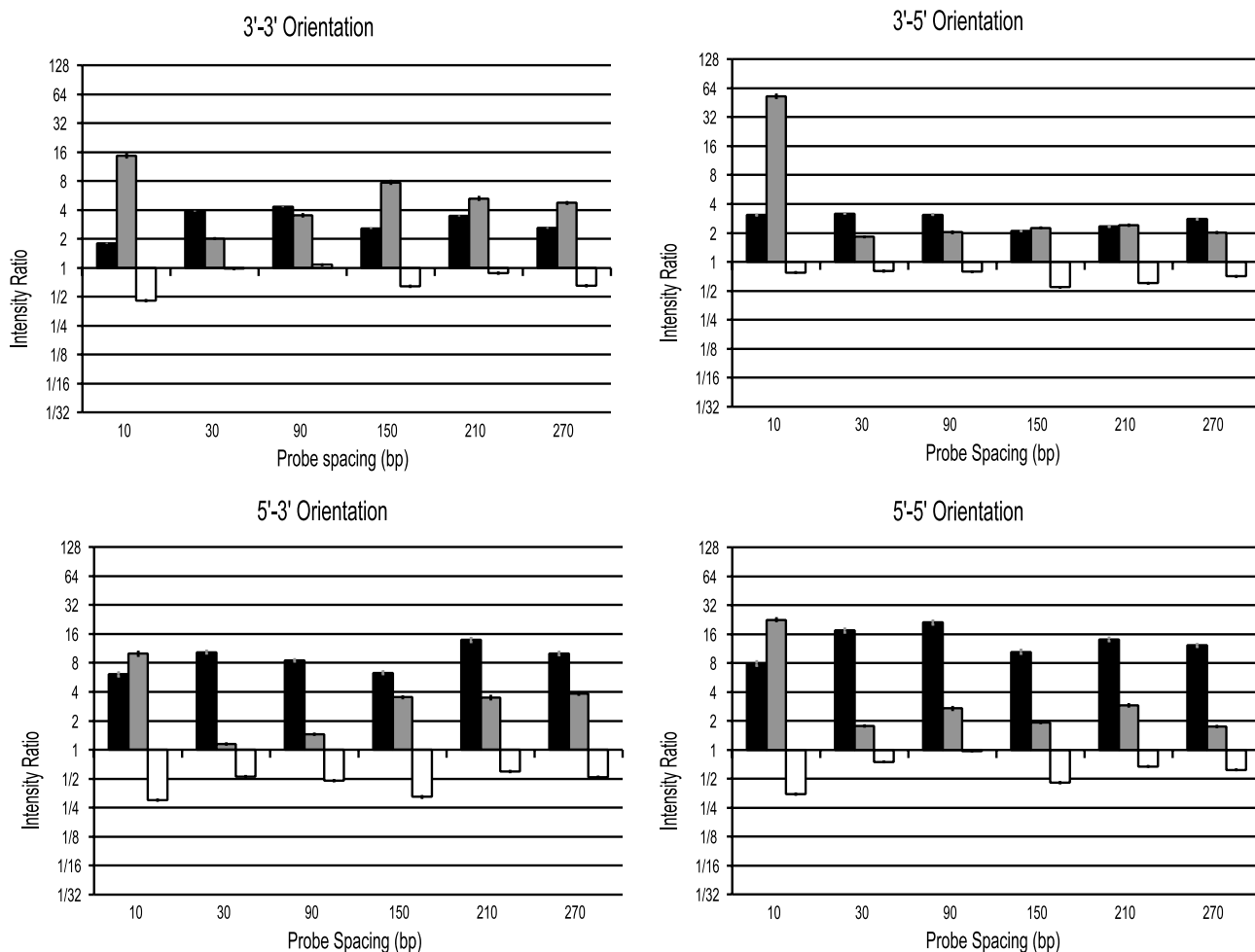


Fig. 3. Hybridization to the N end of the LysA target. Ratios of intensities of mixed probe spots, averaged over six slides, are compared to corresponding single probe spots from the data in Fig. 2. Bars are grouped according to each possible mixed probe pair. Error bars represent 99% confidence intervals. In the data for 5'-3' orientation with variable probe spacing of 30 bases, 2/6 sets of data failed the  $t$  test ( $p > 0.01$ ). All other mixed versus single probe comparisons passed the  $t$  test for significance at  $p < 0.01$ . Black bar,  $I_{\text{mixed}}/I_{\text{fixed}}$ ; dark gray bar,  $I_{\text{mixed}}/I_{\text{variable}}$ ; white bar,  $I_{\text{mixed}}/I_{70\text{mer}}$ .

## Discussion

We have shown that, for a printed microarray, mixed probe spots containing a pair of 30mers show significantly greater hybridization than spots containing a single 30mer and can approach the amount of hybridization to spots containing a 60mer or a 70mer.

Mixed probe spots targeting the C ends did not hybridize as much or as consistently as probes targeting the N ends. Perhaps the target amplification method used here is less efficient for producing the longer Cy3-labeled strands needed to hybridize to the C end probes (Supplemental Figure 7 online). This parallels established strategies for the production of labeled target from total RNA using poly(T) primers: probes need to be targeted to the 5' ends of such bottom strand targets (C ends of the gene) because of incomplete reverse transcription [12] and/or amplification using nucleotide analogs.

In addition to target-end bias, other factors influencing probe selection include avoidance of repetitive

sequences and secondary structure in the target and consistency of estimated  $T_m$  across an entire set of probes. The insensitivity of mixed probe enhancement to probe spacing over shorter distances means that the positions of the two probes in a pair are not dependent on one another. This provides an additional degree of freedom for choosing each pair according to the desired parameters versus choosing a single continuous sequence of twice the length. This increased flexibility results in the optimization of probe selection for mixed probe pairs.

The variable spacing between the probes and the variation in their 5' or 3' attachment to the glass slide was used to seek an optimal configuration of the hybridized target as its chain passes from one probe to the other. As illustrated in Fig. 1A, one can imagine either U- or N-shaped configurations of the target when pairs of probes are attached antiparallel or parallel, respectively. The position of the probes relative to that of the target leaves many hundreds of bases of the target on the 5' or 3' side of the sequence free in solution to interact with other

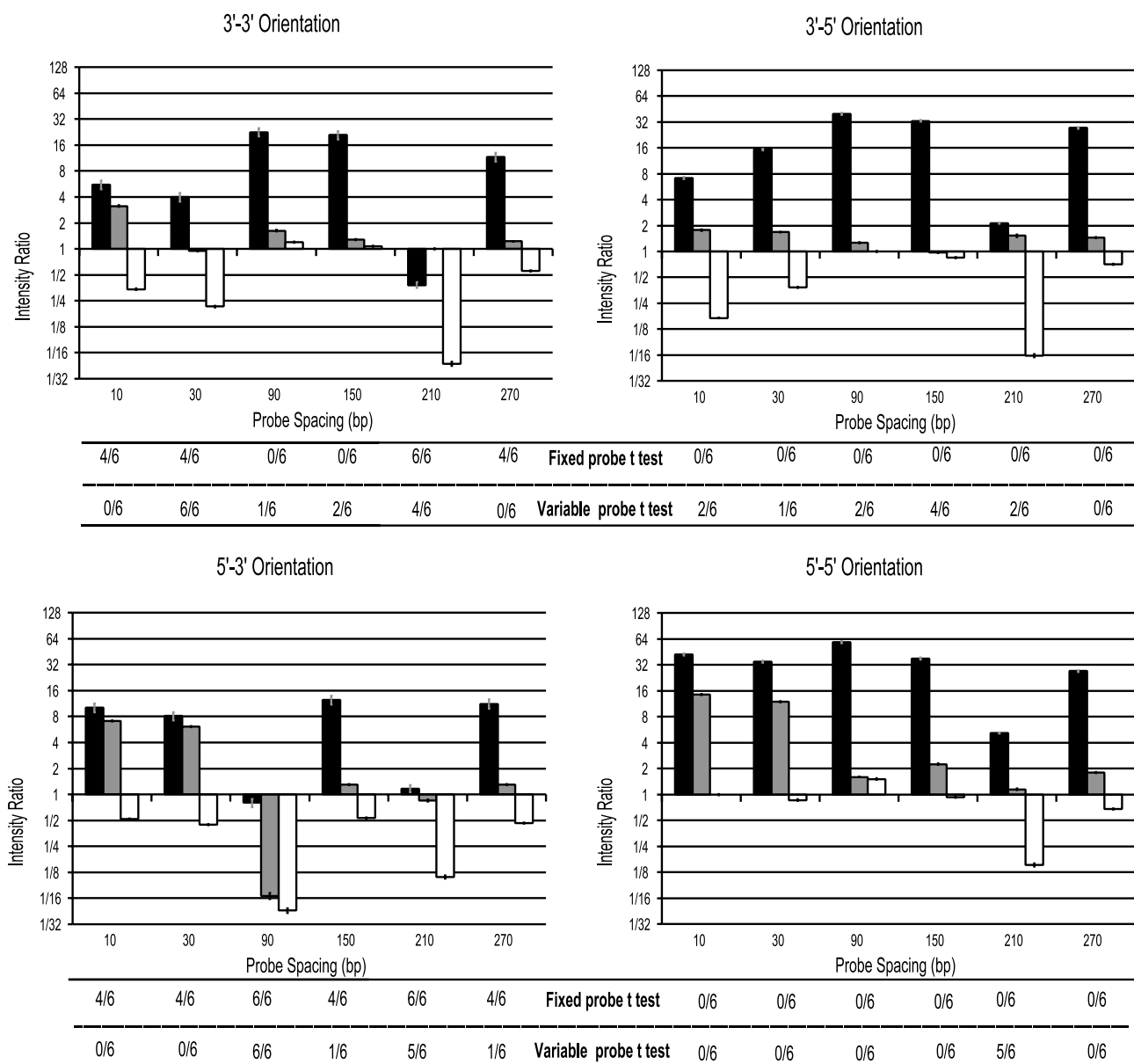


Fig. 4. Hybridization to C end of LysA target averaged over six slides. Ratios of intensities of mixed probe spots are compared to corresponding single probe spots. Bars are grouped according to each possible mixed probe pair. Error bars represent 99% confidence intervals. Tables below each graph give the proportion of slides on which each probe pair failed a  $t$  test ( $p > 0.01$ ) testing whether mixed probe intensity was greater than either the corresponding fixed probe or the corresponding variable probe intensity. Black bar,  $I_{\text{mixed}}/I_{\text{fixed}}$ ; dark gray bar,  $I_{\text{mixed}}/I_{\text{variable}}$ ; white bar,  $I_{\text{mixed}}/I_{70\text{mer}}$ .

target molecules. Our estimate from measuring the attachment of fluorescent 30mer probes suggests that the printed spots have probe spacings of 8–9 nm. This is about four times the closest possible packing of DNA double helices. The situation at the microarray can be compared to that seen in the crowding during feeding by a nursing sow and her litter. Fragmentation reactions, as used in some target preparation protocols, would reduce the length of free target attached to hybridized target but may necessitate closer probe spacings for mixed probe pairs.

The mixed probe technique should be particularly suited to the analysis of alternative splicing [13] and of

single nucleotide polymorphisms [14,15] (SNPs). For instance, if each probe in the pair corresponds to adjacent variant exons in a transcript, the splice variant targets binding to the spot would consist of three populations: unspliced molecules that recognize both probes or spliced target molecules that hybridize to either of the two probes but not both. By comparing these spots to control spots containing each of the two probes paired with a scrambled probe respectively (a scrambled probe being one that does not bind to any splice variant), the relative amounts of splice variants in the sample that contain one or both of the exons can then be determined. The decrease in cooperative signal

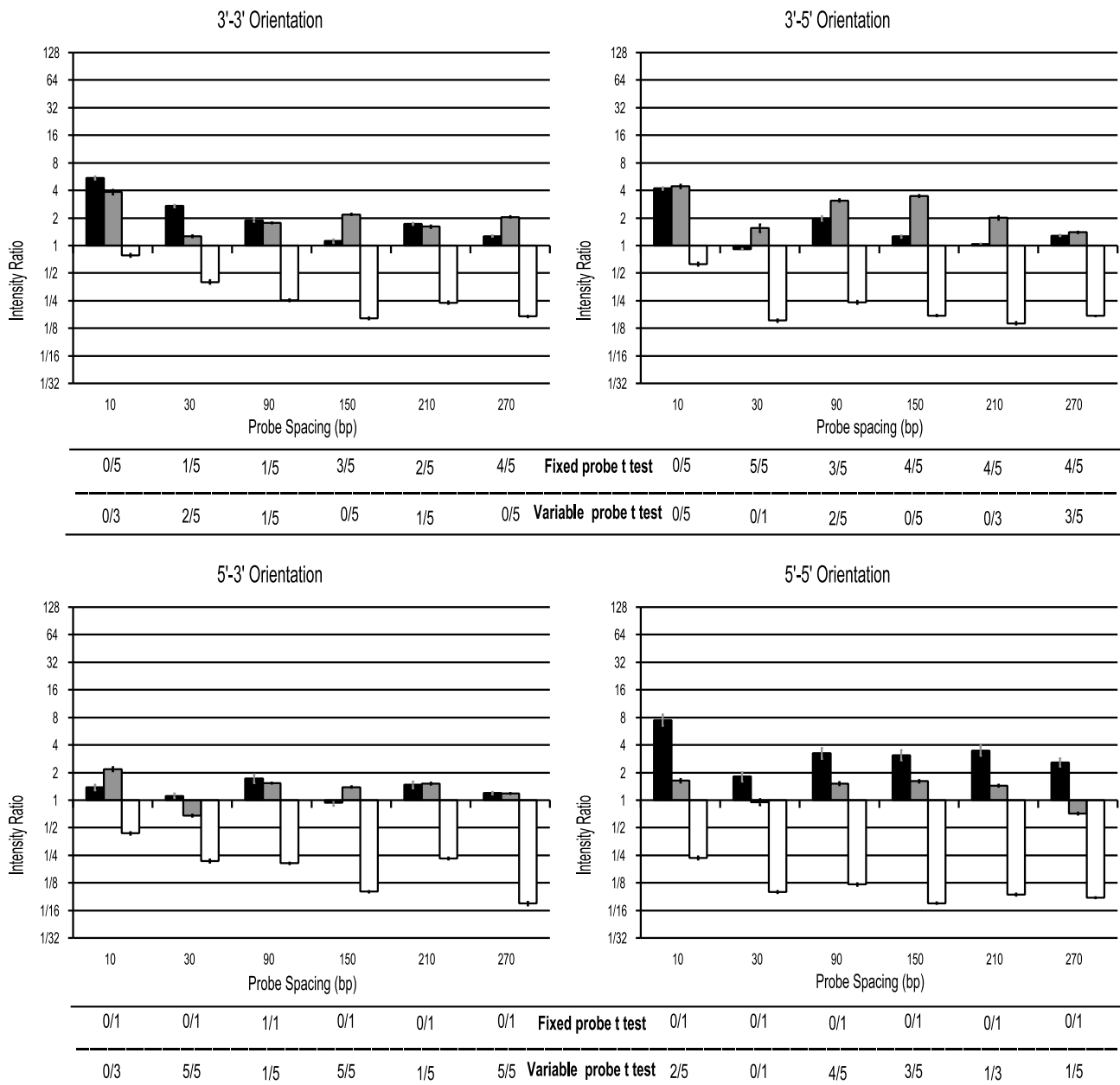


Fig. 5. Hybridization on UltraGAPS slides to the N end of PheB target averaged over five slides. Ratios of intensities of mixed probe spots are compared to corresponding single probe spots. Bars are grouped according to each possible mixed probe pair. Error bars represent 99% confidence intervals. Tables below each graph give the proportion of slides on which each probe pair failed a *t* test ( $p > 0.01$ ) testing whether mixed probe intensity was greater than either the corresponding fixed probe or the corresponding variable probe intensity. Black bar,  $I_{\text{mixed}}/I_{\text{fixed}}$ ; dark gray bar,  $I_{\text{mixed}}/I_{\text{variable}}$ ; white bar,  $I_{\text{mixed}}/I_{60\text{mer}}$ .

when the target spacing is increased beyond 500 bp would be useful for detecting length variants.

Mismatched single probe spots are already used as a control on oligonucleotide arrays [13], and reduced hybridization signal in single base mismatch versus perfect match probes could be useful for detecting SNP sequence variants. A similar signal reduction may occur for mixed half-match spots versus perfect match mixed probe spots. A single base mismatch is enough to partially disrupt binding to a 60mer. Half-match spots where

one of the probes has a single base mismatch would be expected to behave similarly to the half-match spots used here, in which one of the probes is a complete mismatch.

For SNP analysis, the dissociation temperatures of the probe/target double helix with a single base mismatch must be distinguishable from the correct match [16]. The use of shorter probes should offer greater discrimination [16].

Production of mixed probe spot arrays by ink jet or contact printing can be achieved by simply synthesizing



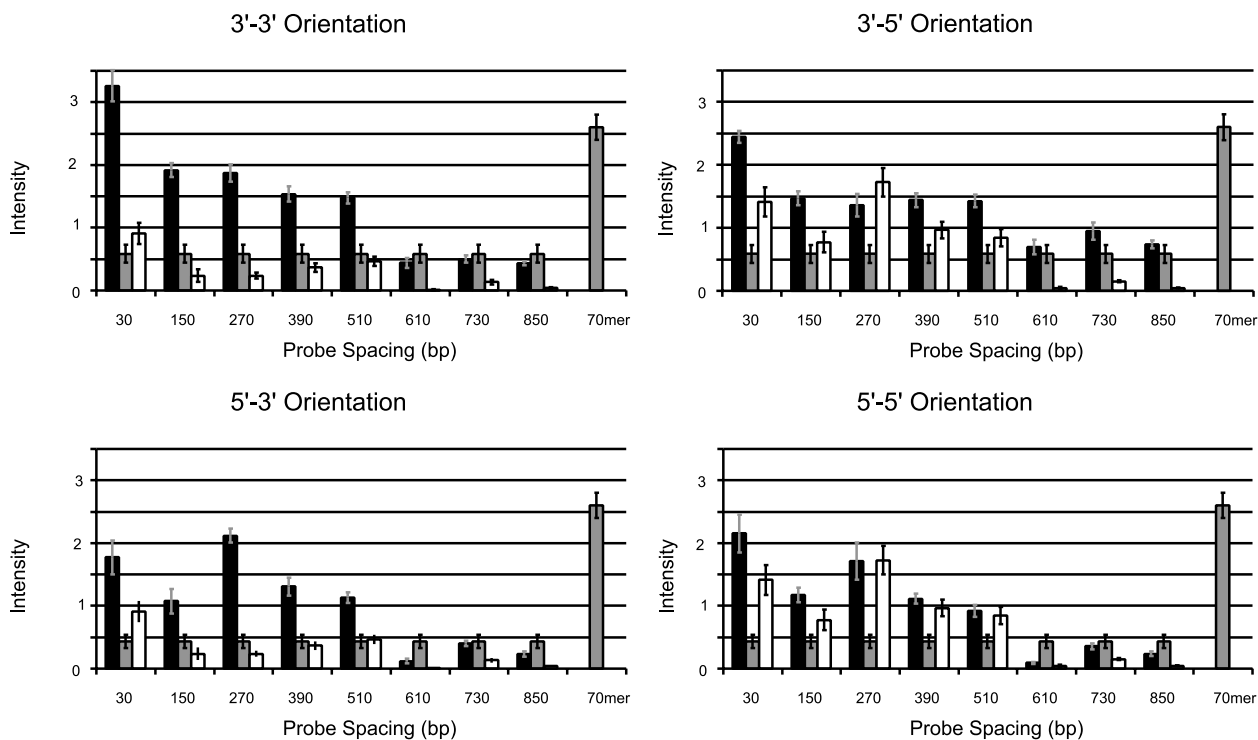


Fig. 6. Hybridization to LysA target with larger probe spacings averaged over three slides. Intensities are in relative fluorescence units ( $10,000\times$ ). Bars are grouped according to each possible mixed probe pair, along with each of the two corresponding single probes. Error bars represent one standard deviation in either direction. Black bar, mixed probes; dark gray bar, fixed probe; white bar, variable probe.

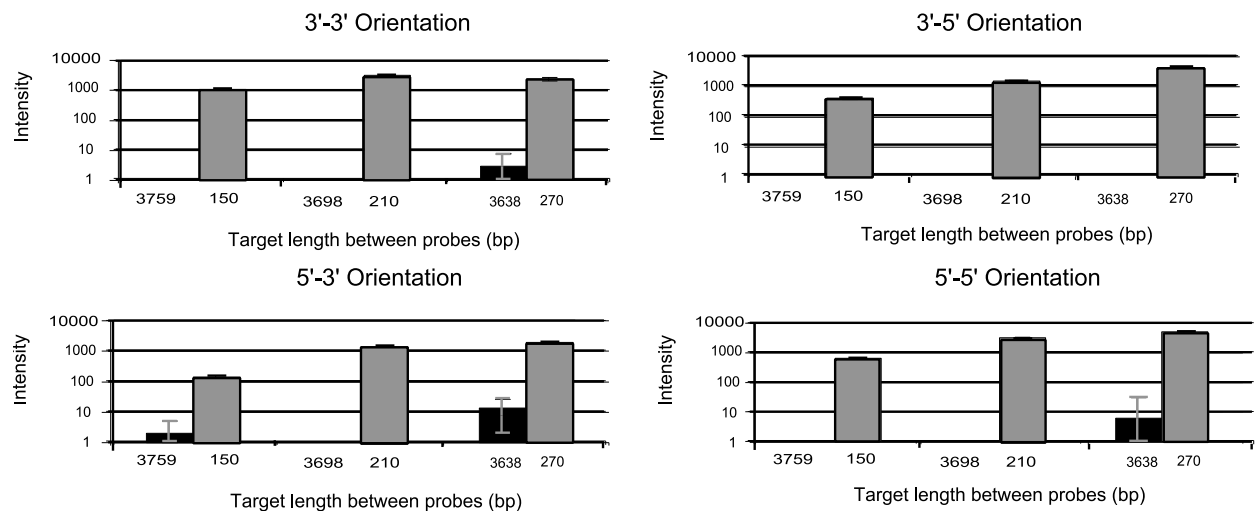


Fig. 7. Hybridization to normal and long variant LysA targets. The ability of the same mixed probe spots to discriminate targets of different lengths is shown as signal for a long variant LysA (3969 bp) averaged over three slides and compared to positive control, normal LysA target. The graphs are shown in logarithmic scale and the intensities are in relative fluorescence units. Bars are grouped according to each possible mixed probe pair. Error bars represent one standard deviation in either direction. Black bar, mixed probes; LysA long variant; dark gray bar, positive control LysA normal target.

oligomers separately and then mixing before deposition. We propose that photolithographic methods could also be used if wavelength-specific deprotection chemistries were available. In theory, two deprotection methods that are specifically sensitive to different wavelengths of light would be sufficient to synthesize separate oligonucleotide strands within the same array spot.

## Conclusion

Mixed probe microarrays have individual spots consisting of two short probes, selected to hybridize at non-contiguous sites in the same targeted polynucleotide. These spots with mixed oligonucleotide probes display cooperative hybridization signals greater than those that can be achieved by either probe alone. Mixed probe

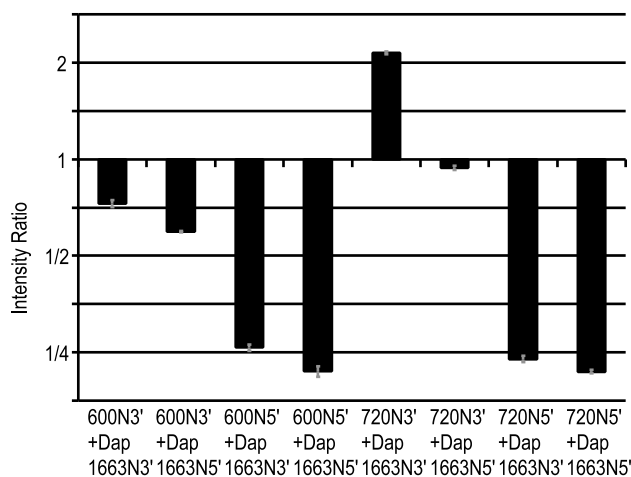


Fig. 8. Hybridization of half-match mixed probe spots to *LysA* target averaged over three slides. Each spot consisted of probe targeting DapB paired with a probe targeting *LysA*. *LysA* N end probes 600 and 700 are the variable probes for the 150- and 270-base spacings, respectively, while Dap1663 is the DapB N end fixed probe. Intensity ratios and 99% confidence error bars are indicated for mixed probes versus the single *LysA* probe.

spots can approach the amount of hybridization to spots containing a 60mer or 70mer. This allows one to take advantage of both the higher synthesis yields of short probes and the greater sensitivity of long oligonucleotides. The mixed probe method results in an increase in the effective probe length without an increase in the actual oligonucleotide length, allowing some of the benefits of long probes to be transferred to photolithographic arrays. Additionally mixed probe microarrays would decrease the cost for printed arrays, while providing additional flexibility. One can avoid inherent targeting problems such as repetitive sequences or secondary structures and it is easier to choose probes with consistent  $T_m$ . This strategy of using mixed probes provides new design options for microarray hybridization assays to detect RNA abundance, RNA splice variants, or sequence polymorphisms.

#### Acknowledgments

We thank A.E. Peritz for helpful discussions and Marie Scearce for technical help. The work was made possible by a grant to D.A.B. and P.L. from the University of Pennsylvania Research Foundation. S.R.B. and A.H. were supported by the Roy and Diana Vagelos Scholars Program.

#### Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at [doi:10.1016/j.ab.2005.03.030](https://doi.org/10.1016/j.ab.2005.03.030).

#### References

- [1] R.J. Lipshutz, S.P.A. Fodor, T.R. Gingeras, D.J. Lockhart, High density synthetic oligonucleotide arrays, *Nat. Genet. Suppl.* 21 (1999) 20–24.
- [2] M. Schena, D. Shalon, R. Davis, P. Brown, Quantitative monitoring of gene expression patterns with a complementary DNA microarray, *Science* 270 (1995) 467–470.
- [3] G.H. McGall, et al., The efficiency of light-directed synthesis of DNA arrays on glass substrates, *J. Am. Chem. Soc.* 119 (1997) 5081–5090.
- [4] M.H. Caruthers, Gene synthesis machines: DNA chemistry and its uses, *Science* 230 (1985) 281–285.
- [5] Z. Guo, et al., Direct fluorescence analysis of genetic polymorphisms by hybridization with oligonucleotide arrays on glass supports, *Nucleic Acids Res.* 22 (1994) 5456–5465.
- [6] T.R. Hughes, et al., Expression profiling using microarrays fabricated by an ink-jet oligonucleotide synthesizer, *Nat. Biotechnol.* 19 (2001) 342–347.
- [7] A.B. Steel, R.L. Levicky, T.M. Herne, M.J. Tarlov, Immobilization of nucleic acids at solid surfaces: effect of oligonucleotide length on layer assembly, *Biophys. J.* 79 (2000) 975–981.
- [8] U.B. Gyllenstein, H.A. Erlich, Generation of single-stranded DNA by the polymerase chain reaction and its application to direct sequencing of the *HLA-DQA* locus, *Proc. Natl. Acad. Sci. USA* 85 (1988) 7652–7656.
- [9] J.P. Brody, B.A. Williams, B.J. Wold, S.R. Quake, Significance and statistical errors in the analysis of DNA microarray data, *Proc. Natl. Acad. Sci. USA* 99 (2002) 12975–12978.
- [10] B. Efron, R. Tibshirani, Statistical data analysis in the computer age, *Science* 253 (1991) 390–395.
- [11] R. Lyne, et al., Whole-genome microarrays of fission yeast: characteristics, accuracy, reproducibility, and processing of array data, *BMC Genom.* 4 (2003) 27.
- [12] E. Talla, F. Tekaiia, L. Brino, B. Dujon, A novel design of whole-genome microarray probes for *Saccharomyces cerevisiae* which minimizes cross-hybridization, *BMC Genom.* 4 (2003) 38.
- [13] G.K. Hu, et al., Predicting splice variant from DNA chip expression data, *Genome Res.* 11 (2001) 1237–1245.
- [14] D.G. Wang, et al., Large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome, *Science* 280 (1998) 1077–1082.
- [15] S.A. Booth, M.A. Drebot, I.E. Martin, L.K. Ng, Design of oligonucleotide arrays to detect point mutations: molecular typing of antibiotic resistant strains of *Neisseria gonorrhoeae* and hantavirus infected deer mice, *Mol. Cell. Probes* 17 (2003) 77–84.
- [16] H. Urakawa, et al., Optimization of single-base-pair mismatch discrimination in oligonucleotide microarrays, *Appl. Environ. Microbiol.* 69 (2003) 2848–2856.