Microarray Probe Design, Validation and Calibration Experiments

David P. Kreil[*†‡] Peter Sykacek[* †‡]
Richard P. Auburn[‡] Lisa A. Meadows[‡]
Roslin R. Russell[‡] Bettina Fischer[‡]
Santiago Sevillano Matilla[‡] Debashis Rana[‡]
Steven Russell[‡] Gos Micklem[‡§¶]

Revision 22nd April 2008

— Supplement —

Enquiries: submission0804@kreil.org

[*] Joint first authorship.
[†] Chair of Bioinformatics, Department of Biotechnology,
Boku University, Vienna,
A-1190 Muthgasse 18, Austria
[‡] Cambridge Systems Biology Centre, University of Cambridge,
Tennis Court Road, Cambridge, CB2 1QR, UK,
& Department of Genetics, University of Cambridge,
Downing Street, Cambridge CB2 3EH, U.K.
[§] Cambridge Computational Biology Institute, Dept of Applied Mathematics and Theoretical Physics, University of Cambridge, Wilberforce Road, Cambridge CB3 0WA, U.K.
[¶] Corresponding author

Required Declarations

The supplementary material which this document references is for confidential preview only. Please treat all data accordingly.

We cannot take responsibility for the availability of material referenced by links to external sites. Please advise us if you encounter any problems.

Main paper abstract

With the growing availability of entire genome sequences, oligonucleotide microarrays have increasingly become an attractive option for the study of gene expression. Careful probe design is important for array performance and potential probe cross-hybridization requires special attention. Implicit in this requirement is the need for careful control of experimental conditions, yet relatively little work has been reported on their systematic examination. Here we introduce a general approach which can be used to optimize any microarray experimental parameter that affects all samples equally. We illustrate this method by designing a widely-used *Drosophila melanogaster* gene expression microarray probe set and then quantifying the impact of different hybridization temperatures on array performance. In the process we were also able to experimentally validate the array design algorithm by showing that probes which were predicted to peform poorly due to cross-hybridization, do perform poorly in practice.

This online supplement is available from http://bioinf.boku.ac.at/pub/indac2008/. Data and scripts are also available from that archive. Please note that the supplement and auxiliary materials are under Copyright © 2006–2008. Additional software is available on request from the authors. (We can either run our code on your data, or collaborate in installing our system at your site.)

Contents

Chapter S-1

Using this archive

S-1.1 Viewing the Supplement and material referenced

This document is provided in PDF format (*cf.* Section S-1.2). Auxiliary information is referenced by HTTP URLs (Hyper Text Transfer Protocol – Universal Resource Locations). If you view this document in a stand-alone browser, *e. g.*, Acrobat Reader, clicking on a link should open a new browser window showing the content to which the link refers.

If you are viewing this document through a plug-in, your browser may loose the original page context when following a link, so when you go back to this document, you might return to the title page. In such a case you may want to save this document to a local disk, and then view it in a stand-alone PDF browser, like Acrobat Reader.

S-1.2 Description of file formats

File formats used in this archive include the following:

- American Standard Code for Information Interchange (ASCII) is used in data files, pre-formatted text for reports, and program code / script files. Columns in data files are typically TAB delimited.
  This format is the simplest and should cause the least problems.

- Adobe Portable Document Format (PDF) for typeset material. This supplement is made available in PDF.

  There are free viewer programs available for this format. To obtain such a viewer, please visit, for example:

  – Ghostscript, Ghostview and GSview from the Computer Sciences Department at the University of Wisconsin-Madison, USA,

  – Adobe Acrobat Reader from Adobe Inc., USA.

  Many browser programs for the World Wide Web can run so-called plug-ins for viewing PDF content.

- Adobe PostScript (PS) for typeset material. To obtain free tools for viewing and printing, please visit, for example, Ghostscript, Ghostview and GSview from the Computer Sciences Department at the University of Wisconsin-Madison, USA. These files are provided for convenience only, and are usually the best format for printing.

- `bzip2` compressed files. Large files (particularly text) may be compressed with `bzip2` for efficiency. Free utilities to unpack such files are available from `http://www.bzip.org/`.

- Microsoft Excel spreadsheet files (XLS) for certain charts and tables.

- Grace plotting tool files (AGR). The original files containing the data to produce some graphs and figures. Although they are text files, you may prefer to view them in the freely available Grace plotting program.

- Various graphics file formats. Typical formats include JPEG, which is a lossy compression format well suited for photos with smooth gradients, and TIFF, which is a particularly flexible format, supporting

both lossy and non-lossy compression schemes (TIFF-FAQ). For viewing or converting many graphics file formats, free tools are available (GraphicsMagick, ImageMagick).

- ZIP archives. Larger collections of files are provided in compressed archives. Free utilities to unpack these archives are available from the Info-ZIP group. Users of the Microsoft Windows system may wish to use WinZIP.

Chapter S-2

Results

S-2.1 Probe design

S-2.1.1 Experiences and results from probe design

An ideal microarray has probes of uniform thermodynamic properties that all sensitively and specifically respond to their respective targets. It is important that targets with similar sequences should be detected with high specificity because transcript variants can have substantially different biological functions. Actual designs may need to make compromises. For example, it might not be possible for probes targetting highly similar transcripts to assess their respective expression levels without strong cross-talk, and one may chose to select one probe assessing these highly similar transcripts together. A set of gene transcript design targets was constructed from annotations of the *D. melanogaster* genome (*Drosophila* Gene Collection v4.0 and *Drosophila* Heterochromatin Genome Project v3.1), including all alternative splicing variants. After addition of 6 markers of common interest and 21 *Arabidopsis thaliana* targets supporting exogenous spikes, targets with sequence similarity of more than 90% were merged, yielding a final target set of 15,829 sequences. We employed OligoArray 2.1 (‘`OA2`’) to generate multiple probe candidates per target transcript. These probe candidates were then

further checked for matches to presumably non-coding genomic sequence and ranked by their thermodynamic properties. Further details are available in the Methods section.

Execution parameters of programs for microarray probe design typically provide threshold criteria for the acceptance of probe candidates. For `OA2`, the probe candidate closest to the 3′ terminus of the target sequence that passes all the following criteria is selected (*8*): probe length and probe–target melting temperature T_m within given ranges, no stable probe secondary structure (self-folding), `GC` content in range (which we did not restrict), no tandem repeats, and a minimal number of predicted stably hybridizing non-target transcripts. Probe length strongly affects the average sensitivity, specificity, noise and bias of probes (reviewed in Kreil *et al.* (*7*)), with probe lengths of 50–150 bases giving a good performance compromise (*4*). Following pilot experiments demonstrating sufficient sensitivity and specificity with the standard protocols employed in our laboratory (*cf.* Methods section) we chose 65–69 base oligonucleotides as probes. Shorter probes, in particular, showed insufficient sensitivity for common applications (data not shown).

When typical probe candidates fall outside the specified design parameter ranges only an unsuitably reduced search space is available for discovering *specific* probes, leading to probe sets with increased cross-hybridization. For our set of target sequences and the given range of permitted probe lengths, thresholds for dependent parameters were optimized for maximal coverage of targets, *i. e.*, minimizing the number of targets with unsuitably restricted search space.

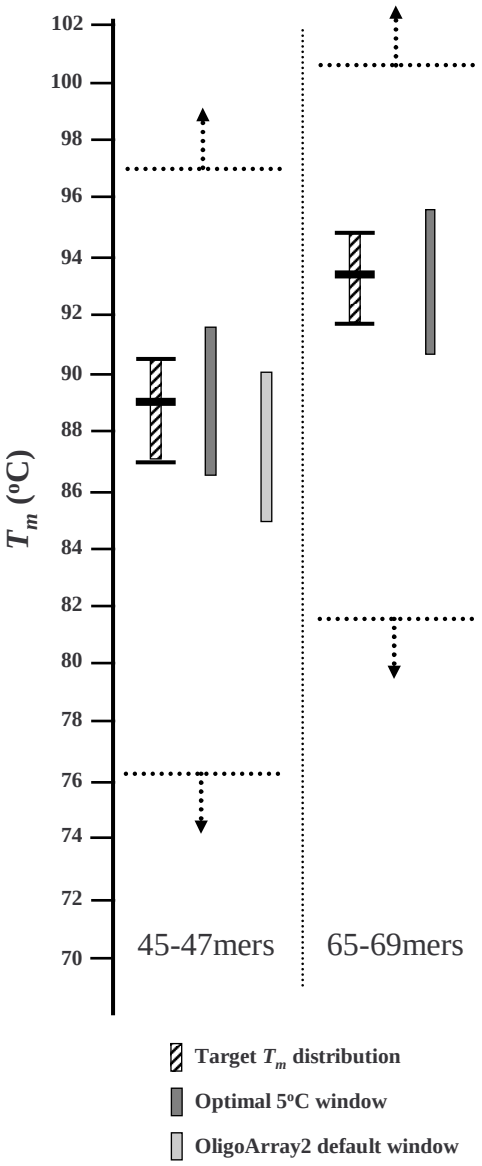

Figure S-2.1: Choice of T_m parameter thresholds to match distributions in probe candidate spaces. Threshold criteria well-matched to the parameter distribution in the probe candidates extend the search space of acceptable candidates, increasing the likelihood that specific probes can be found. The right panel considers 65–69-mers in comparison to 45–47-mer probes (the `OA2` default probe lengths) in the left panel. The distribution of the median T_m of all considered candidate probes per target is shown. The median of these for the set of target transcripts is marked by a black bar; the whiskers indicate first and third quartiles (*cf.* Methods). For the *D. melanogaster* target set in question the `OA2` default T_m window is reasonable for 45–47-mers but needs adjustment for different probe lengths. While a 5°C window could be found to suit most targets, there were targets with unusual properties requiring reruns with different parameter sets. This is illustrated for the most extreme cases: 75% of probe candidates for these transcripts have a T_m beyond the values indicated by the dotted-lines. (Predicted T_m values as calculated by `OA2`.)

An examination of the thermodynamic properties of all available probe candidates led us to adjust design parameters to match the probe properties requested to the naturally observed distributions. Figure S-2.1) shows the distribution of the T_m values of all candidate probes $p \in [1 \dots P_g]$ for all gene transcript targets $g \in [1 \dots G]$, yielding a set of T_m values per sequence, $\mathcal{S}_g = \{T_m^{(g,p)}\}_{\{p\}}$. The dotted lines marking extreme outliers in Fig. S-2.1 delineate $\min\limits_{\{g\}} \left[Q^{(3)}\left(\mathcal{S}_g\right)\right]$ and $\max\limits_{\{g\}} \left[Q^{(1)}\left(\mathcal{S}_g\right)\right]$, where $Q^{(1)}$ and $Q^{(3)}$ denote the first and third quartiles with respect to the set of candidate probes $\{p \mid p \in [1 \dots P_g]\}$. This means that, for the most extreme targets, 75% of all probe candidates for these targets had a melting temperatures beyond the dotted lines. On the other hand, Fig. S-2.1 also illustrates that most targets had melting temperatures in a common range by depicting the distribution of median candidate melting temperatures, median $\mathcal{S}_g$, the median being with respect to the set of candidate probes $\{p\}$. The lower whisker, middle bar, and upper whisker correspond to the first, second, and third quartiles with respect to the target set $\{g \mid g \in [1 \dots G]\}$. For the left panel, $\left(Q^{(1)}, \text{median}, Q^{(3)}\right)\left[\text{median } \mathcal{S}_g\right] = (87.0, 89.0, 90.5)$, whereas the distribution shifts to higher temperatures $(91.7, 93.3, 94.7)$ for longer probes as shown in the right panel.

The `OA2` default parameters for 45–47-mers permit $85°\text{C} \leq T_m \leq 90°\text{C}$, tolerate stable cross-hybridization for melting temperatures $T_x < 65°\text{C}$ and stable probe secondary structure for melting temperatures $T_s < 65°\text{C}$. Examining all possible candidate 45–47-mers, very unusual probe-candidate T_m distributions were observed for some target sequences. As shown in Fig. S-2.1, there were target sequences for which 75% of all probe candidates had melting temperatures less than 76.3°C or for which 75% of all probe candidates had melting temperatures higher than 97°C. Probe design is very difficult for such targets. For any target, the larger the space of probe candidates meeting the initial selection criteria, the more likely it is that a good specific probe with no cross-hybridization potential can be found. We considered a target covered if the probe–target melting temperature T_m was in range for at least 25% of all probe candidates in its 3′-terminal 1500 base pair region.

For a T_m window of given fixed length an optimal location maximizing coverage could be determined by golden section search and successive parabolic interpolation (*3*). By these criteria, the observed range of 45–47-mer melting temperatures was well matched to the default T_m selection window of 85–90°C, for which more than 90% of target sequences were covered. For our target set, the optimal 5°C range maximizing coverage was the interval 86.6–91.6°C, for which 93% of targets were covered.

In comparison, 65–69-mers gave considerably higher melting temperatures, making the default parameters inappropriate. *Less than half* the target sequences were covered using the standard selection window of 85–90°C, reflecting a severely reduced number of candidate sequences that could be considered for probe design. Shifting the default 5°C range by a 5.6°C offset, however, gave a T_m selection window that was well-matched to the natural T_m distribution of probe candidates. Other threshold parameters were adjusted accordingly.

S-2.1.2 Parameters and statistics

There were 54 targets for which no probe could be found (0.3%), the IDs of which are listed in no_probe_found.ids. As can be seen from their sequences in no_probe_found.fasta, they are all particularly short targets, leaving little choice for probe design.

Else, however, there were 15775 targets for which an oligo could be designed (99.7% out of 15829).

Of these, there were

Count	*Predicted cross-hybridization*
14383	targets with a 'perfect' oligo (91.2%)
748	targets with a 'perfect' oligo only failing post-processing (4.7%)
644	targets hitting other genes (4.1%)

Amongst these probes, there were 14467 unique oligo sequences.

Probe oligonucleotide lengths were distributed as follows:

Count	*Probe length*
3173	unique 65mers
712	unique 66mers
733	unique 67mers
749	unique 68mers
9100	unique 69mers

Together with the 7 degradation probes (69mers) this totals to 981663 bases.

Summary statistics of the 48 individual design runs (excl. degradation probes) are shown in Table S-2.1.

Count	*Probe location from target 3′-end*	
14377	within 1500bp	
90	within 2000bp	
Count	*Parameter stringency*	
2286	of highest stringency & standard T_m range	$T_x = T_s = 65°\mathrm{C}$
4169	of high stringency & standard T_m range	$T_x = T_s = 67°\mathrm{C}$
5375	of good stringency & standard T_m range	$T_x = T_s = 69°\mathrm{C}$
298	of good stringency & slightly lower T_m	$T_x = T_s = 67°\mathrm{C}$, $T_m \geq 90.3$
2109	of good stringency & lower T_m	$T_x = T_s = 65°\mathrm{C}$, $T_m \geq 85.0$
230	of good stringency & much lower T_m	$T_x = T_s = 62°\mathrm{C}$, $T_m \geq 81.5$
Count	`OA2` T_m *range*	
6064	requesting a $T_m \in (92.94, 93.44)$	
710	requesting a $T_m \in (92.81, 93.81)$	
1111	requesting a $T_m \in (92.45, 93.95)$	⋮
922	requesting a $T_m \in (92.14, 94.14)$	standard range
1329	requesting a $T_m \in (91.57, 94.57)$	⋮
1602	requesting a $T_m \in (90.65, 95.65)$	
98	requesting a $T_m \geq 90.65$	
298	requesting a $T_m \geq 90.3$	⋮
2109	requesting a $T_m \geq 85.0$	reduced to match
230	requesting a $T_m \geq 81.5$	target composition

Table S-2.1: Relative contributions of design runs with different parameter thresholds. Here, T_m, T_x, and T_s are the probe–target melting temperature, the melting temperature of binding to non-targets, and of self-folded probe structures.

S-2.1.3 Cross-hybridization tables (ΔG)

Systematic all-*vs*-all calculations allowed a comprehensive assessment of predicted cross-hybridization potential. Here, we first consider extremely conservative thresholds supporting the full observed dynamic range of microarray signals (***6***), $R = 10^6$, with a contamination ratio $r < 1\%$, which in the worst case scenario requires that $\Delta G_{\text{diff}} > 12.6\,\text{kcal/mol}$ at $T_{\text{hyb}}^{(\text{eff})} = 70°\text{C}$; see Methods, Eq. (1), in the main manuscript. In order to observe cross-hybridization at this threshold, the contaminating zsequence needs to be in 10^6-fold excess of the target sequence and we must be able to detect intensity deviations as low as 1% of the total signal.

A less extreme but still conservative threshold supports the typical dynamic range of expression intensities seen in a single microarray image scan, $R = 10^3$, with $r < 10\%$ contamination, in the worst case scenario, if $\Delta G_{\text{diff}} > 6.3\,\text{kcal/mol}$. Results were filtered accordingly with each set of thresholds, predicting cross-hybridization potential for 2.5% of probes, with worst-case cross-hybridization detectable in single-scan microarray quantification for only 2.2% of probes. For comparison, predictions for a commercial library were 5.7% and 5.2% of probes, respectively.

In summary, calculations indicated little and weak cross-hybridization as well as good uniformity in hybridization characteristics (Figs S-2.2 and S-2.3) for the newly designed set. The full results of probe design including cross-hybridization matrices by probe and by target are provided below.

Besides the raw unfiltered output, tables were compiled at the two cut-offs described above. Table variants are available to show cross-hybridization by target and by probe for both the novel 'FL002' design and the earlier commercial 'FL001' probe set. There are table variants showing the probes' target IDs ('CG'-number) and variants showing the internally used unique probe IDs. Each row of the tables starts with the target/probe ID and is followed by match groups. Each match group contains the match type ('ok', or 'X' for cross-hybridization), the match ID, and the match ΔG. All fields are tab delimited.

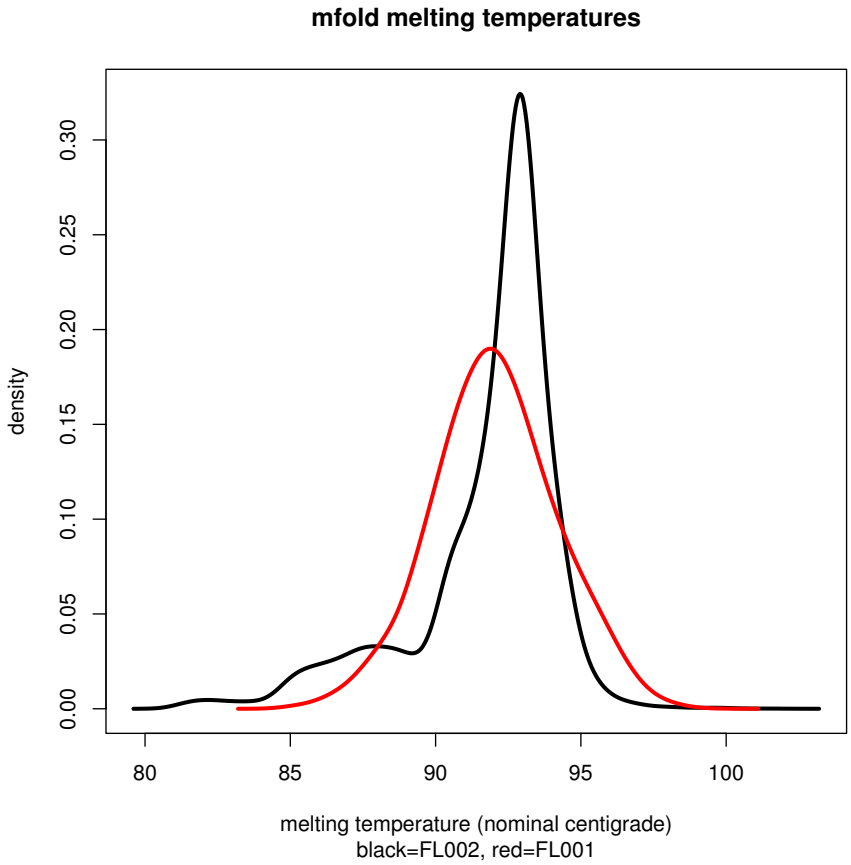

Figure S-2.2: The distribution of probe–target melting temperatures T_m. The T_m distribution is shown for the probe set design described here (FL002; black curve) and an earlier probe set using a commercial library (FL001). Melting temperatures were calculated for full target sequences using `mfold` (*9*). Despite reduced cross-hybridization potential, the new probe set also has a smaller median T_m spread.

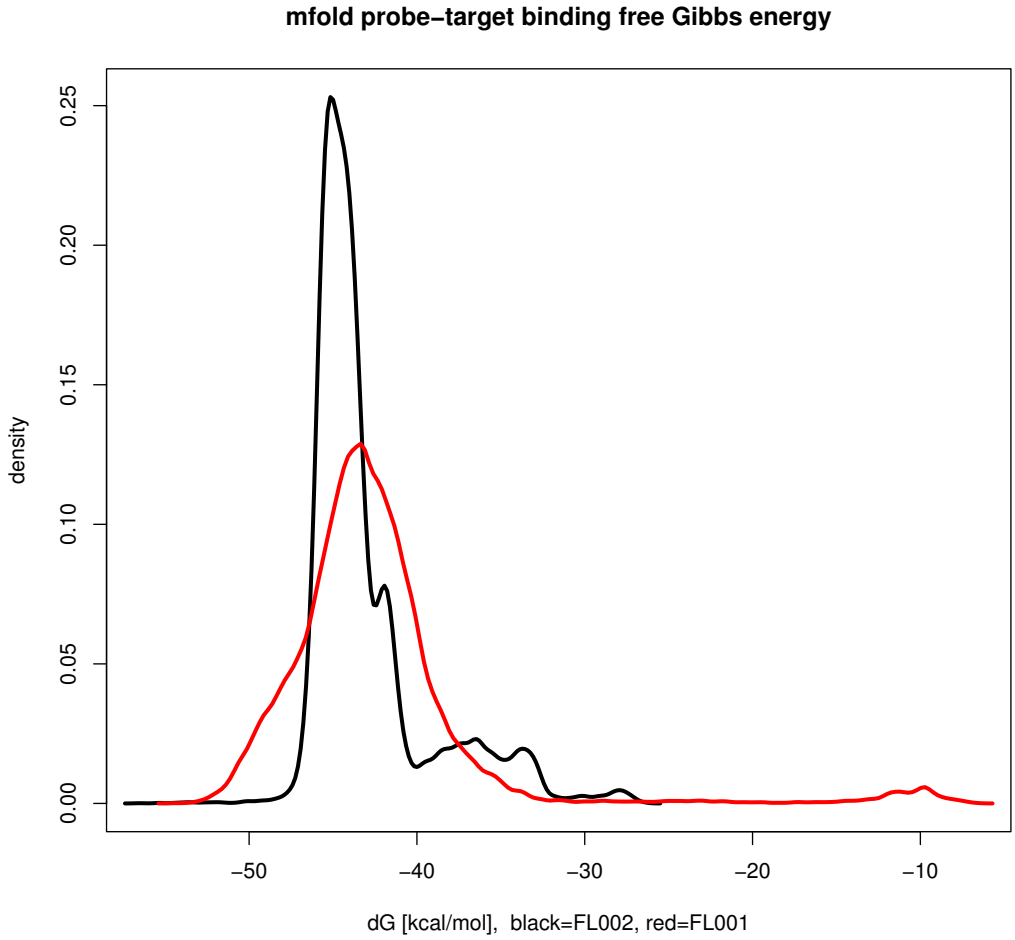

Figure S-2.3: The distribution of probe–target binding free energy ΔG. The ΔG distribution is shown for the probe set design described here (FL002; black curve) and an earlier probe set using a commercial library (FL001). A lower ΔG means higher binding strength. Calculations for hybridization of probes to full target sequences at the design hybridization temperature $T_{\mathrm{hyb}}^{(\mathrm{eff})} = 70°\mathrm{C}$ were performed using `mfold` (***9***). Despite reduced cross-hybridization potential, the new probe set also features a smaller median ΔG spread.

S-2.1.4 Other characteristics

The below figure shows the strong location effect on probe intensity.

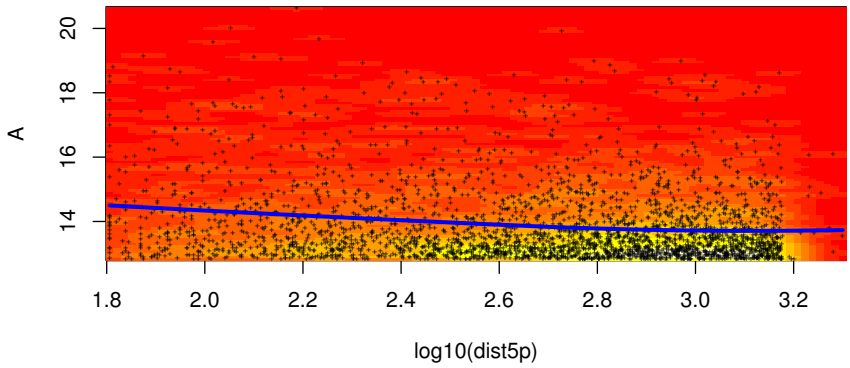

Figure S-2.4: The effect of probe position on average signal intensity. The average Cy3 signal intensity for 4h and 48h hybridizations is shown as a function of the distance of the 5′ terminus of the probe to the 3′ terminus of the target. Targets have been labelled with oligo-dT primers (*cf.* Methods).

S-2.2 Calibration of $T_{\mathbf{hyb}}^{(\mathbf{eff})}$

S-2.2.1 Results overview

Results from the complementary assays described in the paper are collected in Table 1 of the main manuscript. The below figure provides a graphical summary.

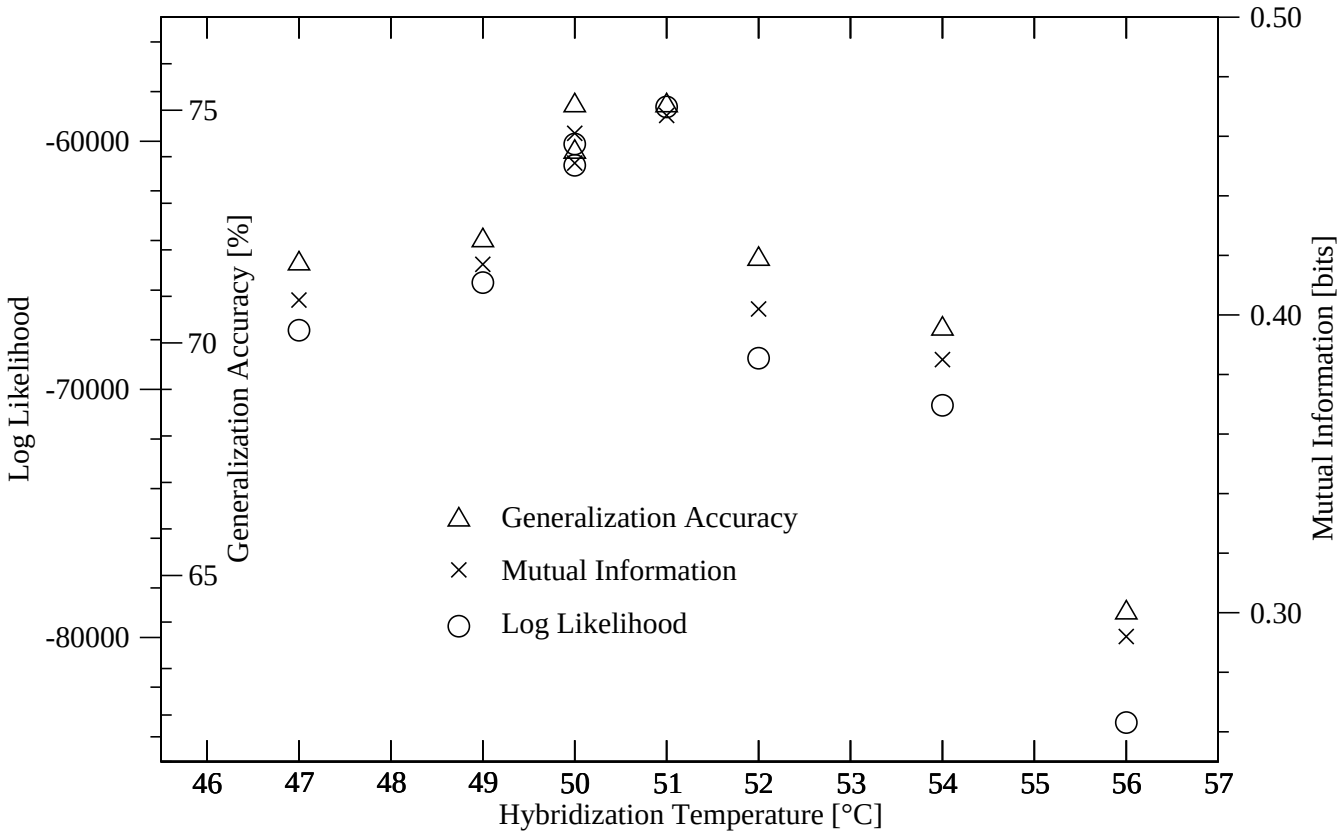

Figure S-2.5: Summary of Performance Measures for Male–Female Separation in response to different hybridization temperatures. This illustration of Table 1 of the main manuscript provides a summary of performance measures for male–female separation.

S-2.2.2 Calculation of generalization accuracy by cross-validation

Complementing these results we provide predictive accuracy estimates based on six-fold cross-validation, in which both channels of one of the six individual slides in turn were used as test samples. All transcripts are used individually to predict whether a channel measurement indicates a male or female sample. For every probe a probit link generalized linear model (GLM) was fitted for the discrimination of male and female samples. The resulting receiver operating characteristic (ROC) curves, estimates of the mutual information, and the generalization accuracies were constructed by pooling the predictions of all probes. A hybridization temperature of 51°C again gave the best predictive performance, as illustrated in the ROC curves (Fig. S-2.6), agreeing with the optimal temperature derived by analysis of the log likelihoods shown in the main manuscript.

S-2.2.3 Effects of cross-hybridization

We employ the ability of probes to discriminate between two distinct biological samples as an indicator of array performance. To be effective, such a measure needs to reflect that cross-hybridization degrades array performance. Both differentially expressed and non-differentially expressed transcripts can cross-hybridize to probes. Cross-hybridization of differentially expressed targets *adds* unwanted information about sample differences to other, non-target probes. In contrast, cross-hybridization of transcripts that are not differentially expressed with probes that should be detecting differential signals will reduce the overall information about sample differences. If the former effect out-weighs the latter, then the overall information about sample differences will be reduced for conditions with increased cross-hybridization, such as lower hybridization temperatures. Consistent with this we have shown that the most informative hybridization temperature was not the lowest one tested.

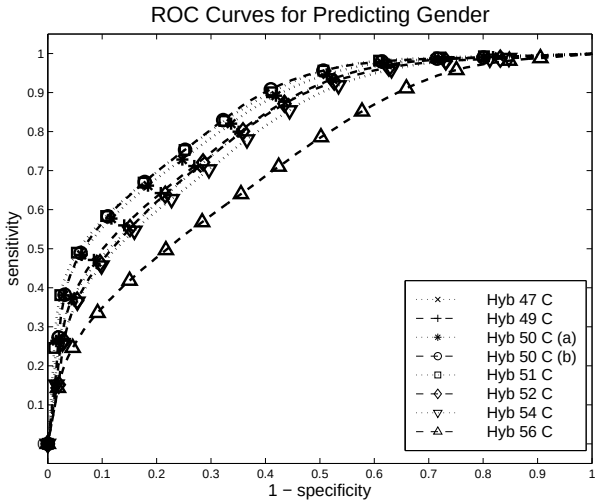

Figure S-2.6: ROC curves for male–female prediction in response to different hybridization temperatures. ROC curves for male–female prediction obtained by six fold cross-testing are shown. All transcripts are used individually to predict whether a channel measurement indicates a male or female sample. We then plot the achieved overall true positive rate (sensitivity) as a function of the accepted overall false positive rate (1 minus specificity). A larger area under the ROC curve corresponds to better performance.

To independently confirm that our results were not affected by cross-hybridization of differentially expressed transcripts we compared the differential expression signal of each probe and that of potential cross-hybridizing non-target probes identified during the probe design process. For every probe with cross-hybridization potential, and for the corresponding non-target probes, we measured the pairwise discrimination performance between target and non-target probes.

The empirical cumulative distribution functions (cdfs) of the log likelihoods (Fig. S-2.7) as well as the receiver operating characteristic (ROC) curves (Fig. S-2.8) both indicate that matching probes are best distinguished from potential cross-hybridizing probes at a physical hybridization temperature of 51°C. This confirmed an absence of bias caused by cross-hybridization of

differentially expressed targets, corroborating the robustness of the original calibration analysis of $T_{\mathrm{hyb}}^{(\mathrm{eff})}$ as reported in the manuscript.

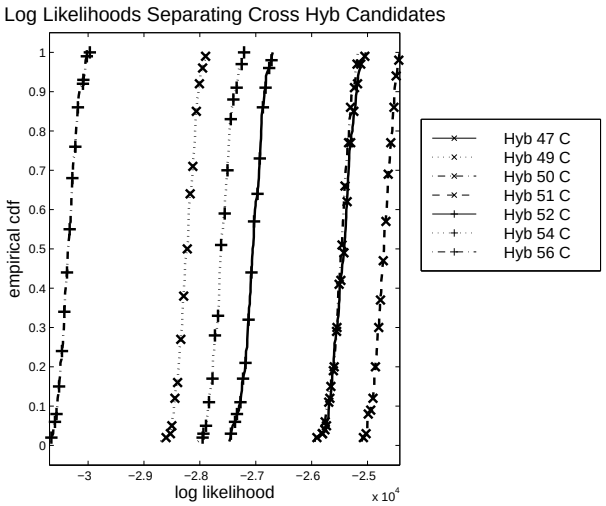

Figure S-2.7: Empirical cdfs of the maximum log likelihood (see main paper for formula) for subsets of 5000 transcripts. The higher the log likelihood, the more evidence the microarray data provide for the sample labels, *i. e.*, the separability of matching probes and potentially cross-hybridizing probes. We again obtained 51°C as optimal hybridization temperature. Also note that the smallest difference between any two experiments is observed for the two independent labelling runs at 50°C indicating a relatively small technical variance.

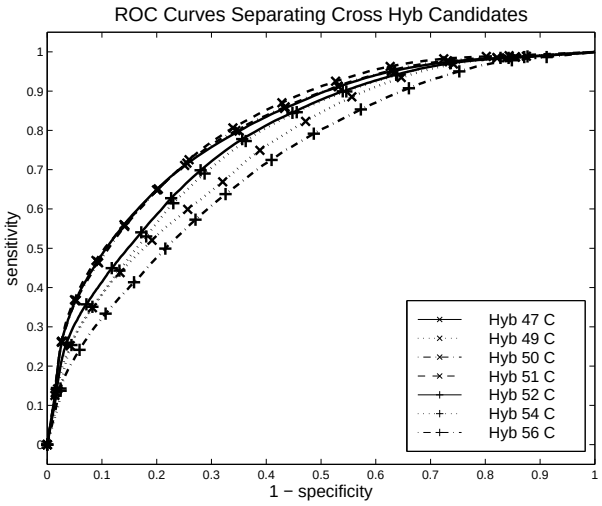

Figure S-2.8: ROC curves for the prediction of a probe being a perfect match or a probe expected to cross-hybridize, obtained by six fold cross testing. All transcripts are used individually for this prediction. We then plot the achieved overall true positive rate (sensitivity) as a function of the accepted overall false positive rate (1 minus specificity). A larger area under the ROC curve corresponds to better performance.

S-2.3 Consideration of hybridization dynamics

With the difficulty of comprehensive spike-in experiments for individual probes, probe properties that can be experimentally assessed in parallel for an entire array probe set, such as probe hybridization dynamics, are of interest. Sets of probes with nonspecific binding have earlier been associated with fast hybridization equilibrium (*5*), and our observations confirm such a general trend.

To examine the effect of hybridization equilibrium in detail, measurements were performed at different hybridization durations of 2, 4, 16, and 48 hours. After confirming that there was indeed a significant and consistent increase of intensity with hybridization duration for most expressed targets as expected, $M(A)$-plots were investigated using the log-ratio $M = s_{48}^{(ch)} - s_4^{(ch)}$ and the corresponding average $A = \frac{1}{2}(s_{48}^{(ch)} + s_4^{(ch)})$ with the raw (not normalized) $\log_2$-scale signal intensities $s_t^{(ch)}$ for hybridization duration t and channel ch (Cy3 or Cy5). Extending the model by Dai *et al.* (*5*), this allowed us to explicitly account for the strong influence that the target concentrations (reflected by A) have on the hybridization dynamics M (*2*). Outliers were detected by constructing a Loess smoother $\overline{M(A)}$ capturing the common trend and a Loess smoother of the standard deviation about this trend $\overline{\sigma_M(A)}$. Probes with $M < \overline{M} - 3\,\overline{\sigma_M}$ were labelled as suspects. The application of a Z-score here corresponds to the `xdev` measure of Dai *et al.* (*5*). We focused on probes with sufficiently strong expression, using the brightest third of signals as arbitrary selection criterion. Results did not change under variation of this selection threshold. For the comparison with the predictions for probe specificity from probe design, not or only weakly-expressed cross-hybridization partners were also removed from the interaction matrices. Results are reported for a first quartile cutoff. Variation of this selection threshold gave no qualitatively different picture.

Hybridization dynamics were examined for all probes by comparing the sig-

nal intensities obtained after different hybridization times. In line with observations relating the speed of reaching equilibrium to hybridization specificity (*5*), probes predicted to cross-hybridize were indeed enriched in sets of probes reaching equilibrium unusually fast (Table S-2.2, Fig. S-2.9).

However, even after taking the effect of target concentrations (*2*) into account, hybridization dynamics could not be used as a predictive indicator of probe performance. Under the population prior, no classification according to design predictions of cross-hybridization could be achieved from hybridization dynamics. Subsampling to equal priors also only gave a classification accuracy of 52%.

Probes reaching equilibrium unusually fast also did not display reduced sample separation performance (data not shown). In contrast, sets of probes predicted to cross-hybridize in the probe design process did show a much reduced sample separation performance as expected (main manuscript, Fig. 2), strongly supporting that prediction.

The large observed scatter suggests further factors confounding the effect of probe specificity on hybridization dynamics. The difficulty of exploiting such indirect measures of probe quality predictively is exacerbated by the fact that in a well-designed array for gene expression profiling only a minority of probes will show nonspecific hybridization. This is also reflected in the small number of probes showing unusual hybridization dynamics on our array.

Dye	fast	normal or slow	Fisher's exact test
Cy3	21 x, 7 o, 89 +	179 x, 174 o, 4342 +	$P_x = 5 \times 10^{-9}$,
	(117 total)	(4695 total)	$P_{x,o} = 6 \times 10^{-8}$
Cy5	18 x, 1 o, 76 +	180 x, 168 o, 4369 +	$P_x = 5 \times 10^{-8}$,
	(95 total)	(4717 total)	$P_{x,o} = 7 \times 10^{-5}$

Table S-2.2: Enrichment of probes predicted to cross-hybridize in the set of probes showing fast hybridization dynamics. Enrichment of probes predicted to cross-hybridize (to non-target mRNA 'x' or to genomic DNA 'o') over specific probes ('+') in the set of probes showing fast hybridization dynamics. Signals have been averaged over two slide pairs.

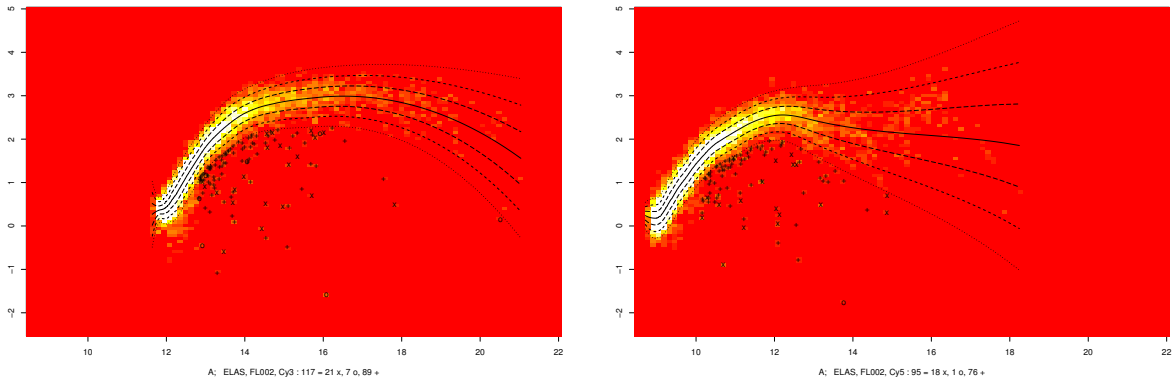

Figure S-2.9: Hybridization dynamics $M(A)$ density maps. Hybridization dynamics $M(A)$ plots, with $M = s_{48}^{(ch)} - s_4^{(ch)}$ and the corresponding average $A = \frac{1}{2}(s_{48}^{(ch)} + s_4^{(ch)})$ for the raw (not normalized) $\log_2$-scale signal intensities $s_t^{(ch)}$ of channel ch (Cy3 or Cy5) at hybridization duration t. A Loess smoother capturing the common trend is shown (solid line). One, two, and three standard deviations about this trend are shown as long dashed, dashed, and dotted lines. Probes with M below the lower dotted line were considered 'fast'. For the brightest third of signals, fast probes are labelled according to their predicted cross-hybridization behaviour: specific '+', matching genomic DNA 'o', or matching non-target mRNAs 'x'. The heat map reflects the densities of transcripts in a region of the plot.

S-2.4 Independent assays: Comparative Genomic Hybridization

As an independent approach for validating array quality, we report the results of an analysis of eight microarrays used for Comparative Genomic Hybridizations (CGH). See the Methods section for a detailed description of CGH experiments and laboratory protocols.

While the observed scatter does not permit probe-level resolution, regions of deletion and duplication can easily be detected as shown by the local average, drawn as a thick line (Fig. S-2.10). The local average was computed by a lowess smoother (local linear regression, span set to 1).

S-2.4.1 Raw data and scripts

The data files and scripts used and additional diagnostic plots are provided. These also show that results are independent of normalization choice.

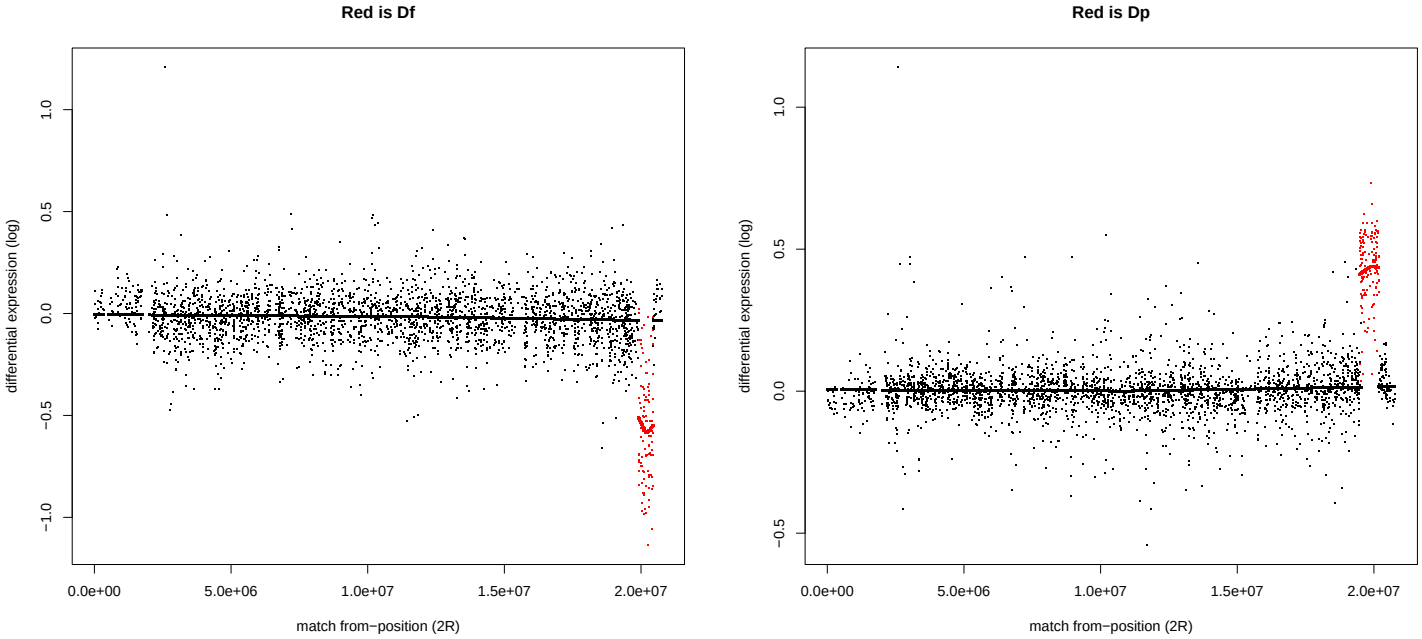

Figure S-2.10: The x-axis represents the probe location along chromosome arm 2R, the y-axis shows relative binding strength. Each dot plots a probe signal relative to the wild type reference (comparable to a log-ratio). The left panel contains results for a genomic deletion (Df), the right panel for a genomic duplication (Dp), the locations of which are shown in red. The thick lines indicate local averages.

Chapter S-3

Methods

This chapter contains the detailed protocols. Additional information is also available on www.flychip.org.uk.

S-3.1 Probe sequence design, technical issues

Here we briefly outline some practical considerations in probe sequence design that were of technical and/or transient nature.

S-3.1.1 Construction of the target transcript set

In making a set of sequences non-redundant, the headers (names) of redundant sequences are usually merged, leading to very long sequence headers. This can at times trigger a malfunction in the BLAST program. The latest version of BLAST (*1*) available at that time was employed (v2.2.10) but the problem also affected earlier available versions. The cause for this malfunction was not further investigated. The names of all target sequences therefore had to be reduced to unique headers not longer than 60 characters.

S-3.1.2 Employment

Minor changes to the OligoArray 2.1 source code allowed us to work around a disruptive `BLAST` output bug. This fix will be made available in the next version of OligoArray to be released later this year (J.-M. Rouillard, pers. comm., 2005). We also adapted the sources to allow multiple instances of the program to share a working directory which simplified distributed dispatch.

S-3.1.3 Raw data and scripts

The data files and scripts used are provided.

S-3.2 Probe spotting

The probe layout used is available online for FL001 (the Operon set) and FL002 (the newly designed set).

S-3.3 Slide processing

Slides from commercial sources were processed according to protocols recommended by the manufacturer.

S-3.4 Hybridization Protocol

All hybridizations and washes were performed using an automated hybridization station (Genomic Solutions GeneTAC HybStation).

Sample preparation, labelling

100 μg of total RNA was extracted from male or female *D. melanogaster* following the group's protocol optimized for large scale extraction of RNA from adult flies. RNA quality was verified by electrophoresis and ethidium bromide staining as well as UV spectrometry using an in-house calibrated Nanodrop ND-1000 spectrophotometer. RNA was then labelled by direct incorporation of Cy3-dCTP (Amersham, Cat. No. PA 53021) or Cy5-dCTP (Amersham, Cat. No. PA 55021) in a reverse transcription reaction primed by anchored oligo $(dT)_{23}$ (Sigma, Cat. No. 04387) using Superscript III Reverse Transcriptase (Invitrogen, Cat. No. 18080-044). This was repeated twelve times for each combination of dye and gender, giving a total of $12 \times 2 \times 2 \times 100\,\mu$g of labelled RNA. All male-Cy3 / female-Cy5 samples were then pooled and split into 12 aliquots, sufficient for 24 arrays. Similarly, all female-Cy3 / male-Cy5 samples were pooled and aliquoted. Aliquots were the dried down with a speed vacuum and stored at $-20°$C. Full details of the standard labelling protocol of the group that was used are available online.

Before hybridization, the required number of labelled sample aliquots were resuspended in Ocimum hybridization buffer (Biosolutions, Cat. No. 1180-200000) and sonicated salmon sperm DNA equivalent to 20ug per array (Invitrogen; Cat. No. 15632-011), pooled, and split into aliquots corresponding to the number of arrays to hybridize.

Hybridization and washes

Hybridization was performed with an automated GeneTAC Hybridization Station (Genomic Solutions). Standard post hybridization washes were performed manually. The group's standard hybridization protocol was employed, with hybridization temperature and duration varied as described in the Methods section of the main manuscript.

CGH experiments and protocols

Four replicates (two plus two dye-swap) were each employed to compare genomc DNA (gDNA) with either a genomic deletion (Df) or a genomic duplication (Dp) to wild type gDNA. A detailed description of the employed strains, including their genomic deletion and duplication, as well as specifics for extracting and labelling genomic DNA (gDNA) are provided in an auxiliary CGH protocol.

S-3.5 Data acquisition and post-processing

Arrays were scanned using a GenePix 4000B dual laser scanner and GenePix Pro 5.1 imaging software (Axon Instruments). Arrays were scanned at 5 μm resolution, simultaneously in both the Cy3 channel (excited by a 532nm laser) and the Cy5 channel (excited by a 635nm laser). Laser power was always set at 100% but photomultiplier tube (PMT) gain was separately adjusted for each channel in order to balance the signal from the two channels and to scan at the highest PMT value for which there were not more than a handful of saturated spots. For the assessment of hybridization dynamics, slides were subsequently scanned again at higher PMT gain settings. Data from multiple scans was then combined by extrapolation using a robust linear model (`rlm` of `R` with default parameters; unpublished).

S-3.5.1 Raw data and scripts

The data files and scripts used are provided.

Application	*Cy3*	*Cy5*
Calibration and validation	550	600
Hybridization dynamics 4/16/48h	530, 730	620, 820
Hybridization dynamics 2/4/48h	500, 640, 1000	580, 720, 1000

Table S-3.1: PMT gain employed (on a scale 0 . . . 1000)

Bibliography

1. Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**(17), 3389–3402. Cited on page: 25

2. Bhanot, G., Louzoun, Y., Zhu, J., and DeLisi, C. (2003) The importance of thermodynamic equilibrium for high throughput gene expression arrays. *Biophys. J.* **84**(1), 124–135. Cited on page: 20, 21

3. Brent, R. P. (1973) *Algorithms for Minimization without Derivatives.* Prentice-Hall, Englewood Cliffs, New Jersey, U. S. A.. Cited on page: 8

4. 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**(12), e99. Cited on page: 5

5. Dai, H., Meyer, M., Stepaniants, S., Ziman, M., and Stoughton, R. (2002) Use of hybridization kinetics for differentiating specific from non-specific binding to oligonucleotide microarrays. *Nucleic Acids Res.* **30**(16), e86. Cited on page: 20, 21

6. Dudley, A. M., Aach, J., Steffen, M. A., and Church, G. M. (2002) Measuring absolute expression with microarrays with a calibrated reference sample and an extended signal intensity range. *Proc. Natl. Acad. Sci. U. S. A.* **99**(11), 7554–7559. Cited on page: 11

7. Kreil, D. P., Russell, R. R., and Russell, S. (2006) Microarray oligonucleotide probes. *Methods Enzymol.* **410**, 73. Cited on page: 5

8. Rouillard, J. M., Zuker, M., and Gulari, E. (2003) OligoArray 2.0: design of oligonucleotide probes for DNA microarrays using a thermodynamic approach. *Nucleic Acids Res.* **31**(12), 3057–3062. Cited on page: 5

9. Zuker, M. (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* **31**(13), 3406–3415. Cited on page: 12, 13