

External RNA Controls for Monitoring Performance of CodeLink™ Whole Genome Bioarrays

As part of the CodeLink™ Expression Bioarray System, CodeLink™ Whole Genome bioarrays provide comprehensive expression profiling on a single bioarray.

CodeLink™ Whole Genome bioarrays include a new set of positive control probes designed against six *E. coli* genes, and when used in conjunction with the mRNA spikes currently supplied with CodeLink™ Expression Assay Reagent Kit, can be used to assess target labeling and assay hybridization processes.

Here we describe two ways in which the CodeLink™ Whole Genome control probes can be utilized in determining array assay quality. In the first method, individually labeled cRNAs corresponding to the bacterial probes can be spiked into complex background cRNA samples. Results obtained using this method more accurately reflect bioarray performance because they are independent of RNA starting material and purity; this robustness makes the method ideal for assessing a within-slide dynamic range. In the second method, each individual bacterial mRNA is spiked at different concentrations into either total or polyA(+) mRNA samples. The signal intensities obtained for each spike can then be measured, providing a way to assess target preparation efficacy in conjunction with overall microarray performance at each transcript level.

In addition to providing dilutions and spiking procedures for both methods, this application note describes the use of CodeLink™ Whole Genome control probes in evaluating bioarray sensitivity and dynamic range.

Products Used

CodeLink™ Human Whole Genome Bioarray	300026
CodeLink™ Expression Assay Reagent Kit (24 reactions)	320012
CodeLink™ Expression Analysis v4.0 software	310030
Cy™5-Streptavidin fluorophore conjugate	PA45001

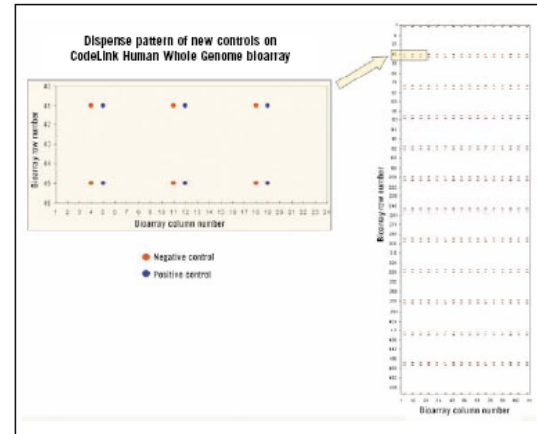


Fig. 1. Dispense layout of positive and negative controls on CodeLink™ Human Whole Genome bioarrays. The bioarray layout is represented by row number on the y-axis and column number on the x-axis. Positive (blue) and negative (red) controls are dispensed next to each other on the array as illustrated in the close-up view on the left (highlighted in yellow). This dispense pattern enables localized comparisons of signal strength between noise (negative controls) and spiked transcripts (positive controls). Note the even dispersion of probes across the entire bioarray surface.

Other Products Required

- GenePix™ 4000B array scanner (Axon Instruments)
- 10 mM Biotin-11-UTP (PerkinElmer)
- Q1Aquick™ PCR purification kit (Qiagen)
- RNeasy™ Mini Kit (Qiagen)

Protocol

1. Method for spiking bacterial transcripts at the cRNA level prior to hybridization

1.1 Prepare cRNA from each individual bacterial mRNA by following the CodeLink™ target preparation protocol (1). Start with 200 ng of each bacterial mRNA in separate tubes for each of the six bacterial mRNAs (*araB*, *entF*, *fixB*, *gnd*, *hisB*, *leuB*).

Note: Biotin will be incorporated into each cRNA during *in vitro* transcription. Enough cRNA will be produced from each bacterial mRNA for thousands of bioarray experiments. Store cRNA samples in small aliquots at -70 °C for up to six months.

1.2 Dilute the cRNA samples to a 0.1 µg/µl concentration.

Note: It is recommended to run cRNA samples on an Agilent 2100 Bioanalyzer to assess integrity. The distribution of cRNA size should be centralized near 1000 bases for *araB*, *fixB*, *hisB*, and *leuB*, while *entF* and *gnd* will be near 1300 bases in length.

1.3 Prepare the cRNA spike-dilution mix by making a 10-fold dilution of each bacterial cRNA.

Note: After making the 10-fold dilutions of cRNAs, the concentration of each spike will be 0.01 µg/µl. We do not recommend storing cRNA at low concentrations for extended periods. Only prepare enough diluted cRNA as necessary for a particular series of experiments.

1.4 Add the following volumes of cRNA and nuclease-free water to six separate tubes.

Note: Different volumes of each spike are added below to normalize each transcript to approximately 1000 bases in length.

- Tube 1: 6.2 µl *araB* + 3.8 µl H₂O
- Tube 2: 7.5 µl *entF* + 32.5 µl H₂O
- Tube 3: 6.0 µl *fixB* + 154.0 µl H₂O
- Tube 4: 8.2 µl *gnd* + 631.8 µl H₂O
- Tube 5: 6.2 µl *hisB* + 633.8 µl H₂O
- Tube 6: 6.4 µl *leuB* + 633.6 µl H₂O

1.5 Make a four-fold dilution of *hisB* by adding 10 µl of the solution from Tube 5 to 30 µl nuclease-free water. Label the new tube "5D" (4x dilution of original Tube 5).

1.6 Make a 16-fold dilution of *leuB* by adding 10 µl of the solution from Tube 6 to 150 µl nuclease-free water. Label this new tube "6D" (16x dilution of original Tube 6).

1.7 Create the final spike-dilution mix by adding 10 µl of Tubes 2, 3, 4, 5D, and 6D to Tube 1. The total

volume of this solution will be 60 µl, and its final concentration will be 1.0 ng/µl.

Note: The final 1.0 ng/µl concentration of the spike-dilution mix is approximately 1,000 fold lower in concentration relative to a typical cRNA target preparation of 1.0 µg/µl. This 1,000 fold concentration difference from spike mix to complex background cRNA simplifies the creation of a within-slide dynamic range at different levels. For example, equal volumes of spike-dilution solution and tissue cRNA (if 1.0 µg/µl) will create a within-slide dynamic range starting at the highest mass ratio of 1:1,000 down to 1:1,024,000 relative to complex cRNA background. However, spike no higher than a mass ratio of 1:2,000 to avoid saturation. Adding the spike-dilution mix at a 1:2,000 mass ratio relative to complex cRNA will produce a dynamic range from approximately 50 fM to 50 pM. This is the suggested spiking range. Results from following this protocol are represented later in this paper.

To add the spike-dilution mix to 10 µg of cRNA at a concentration of 1.0 µg/µl, first add 10 µl of cRNA, followed by 5 µl of spike-dilution solution, and then continue with fragmentation and subsequent hybridization.

Example: Adding the spike-dilution mix in a 1:2,000 mass ratio to complex cRNA will generate the following final mass ratios and hybridization molarities, under standard hybridization conditions (2):

Spiked bacterial transcript	Final mass ratio	Final concentration (pM)
<i>araB</i>	1:2000	51.20
<i>entf</i>	1:8000	12.80
<i>fixB</i>	1:32 000	3.20
<i>gnd</i>	1:128 000	0.80
<i>hisB</i>	1:512 000	0.20
<i>leuB</i>	1:2 048 000	0.05

2. Method for spiking bacterial mRNA transcripts into total RNA

Note: For a typical mammalian cell, mRNA comprises 1-5% of the total RNA content. An estimate of the mass ratios of the bacterial control mRNAs to the target mRNA within the total RNA can be made by assuming 5% mRNA content (3). However, since mRNA abundance can vary according to cell type, spiking at the cRNA level (see prior method) will more consistently yield an accurate assessment of bioarray sensitivity.

2.1 Thaw the bacterial control stock mRNAs on ice and aliquot the following into six individual tubes, adding nuclease free water as noted:

- Tube 1: 4 μ l *araB* (0.1 μ g/ μ l)
- Tube 2: 4 μ l *entF* (0.1 μ g/ μ l) + 6 μ l H₂O
- Tube 3: 4 μ l *fixB* (0.1 μ g/ μ l) + 36 μ l H₂O
- Tube 4: 4 μ l *gnd* (0.1 μ g/ μ l) + 156 μ l H₂O
- Tube 5: 4 μ l *hisB* (0.1 μ g/ μ l) + 636 μ l H₂O
- Tube 6: 1 μ l *leuB* (0.1 μ g/ μ l) + 636 μ l H₂O

2.2 Mix by tapping the sides of the tubes, then centrifuge for 5 s at $\geq 10\,000 \times g$ to gather the liquid. Store the tube on ice.

2.3 Prepare the spike-dilution mix by adding the following to Tube 1:

- 4 μ l *entF* (25.00 ng/ μ l)
- 4 μ l *fixB* (6.25 ng/ μ l)
- 4 μ l *gnd* (1.56 ng/ μ l)
- 4 μ l *hisB* (390/60 pg/ μ l)
- 4 μ l *leuB* (97.70 pg/ μ l)
- 36 μ l nuclease-free water

Note: The final volume in Tube 1 will be 60 μ l.

2.4 Mix well by tapping the side of the tube and centrifuge for 5 s at $\geq 10\,000 \times g$ to gather the liquid. Store the tube on ice.

Note: The resulting concentration of this spike-dilution mix will be 10 ng/ μ l. This RNA solution can be stored at -70 °C for up to one month. We recommend storing this solution in smaller aliquots to eliminate multiple freeze/thawing upon consecutive use.

2.5 The spike-dilution mix can now be added to total RNA samples in a 1:20 000 mass ratio or to polyA(+) mRNA in a 1:1000 mass ratio. Adding the spike-dilution mix at these recommended mass ratios relative to total and mRNA will yield the following individual mass-ratio levels:

<i>araB</i>	1:1000
<i>entF</i>	1:4000
<i>fixB</i>	1:16 000
<i>gnd</i>	1:64 000
<i>hisB</i>	1:256 000
<i>leuB</i>	1:1 024 000

Note: When spiking into total RNA, the mass ratios above assume 5% of the total RNA represents polyA(+) mRNA with an average transcript length of 1000 bases. The only assumption when spiking into mRNA is that the average mRNA transcript length is 1000 bases. Average mRNA length can be determined for mRNA samples using an Agilent 2100 Bioanalyzer with 1 kB RNA ladder. This type of length verification cannot be performed on total RNA due to the abundance of ribosomal RNAs. Once you have determined the average mRNA length for a desired sample, you can adjust the level of mass ratio spiking.

Examples:

- For 1 μ g of total RNA, add 50 pg of spike-dilution mix (at a concentration of 10 ng/ μ l). This represents a 1:20 000 mass ratio of spike to total RNA sample. Proceed by:
 1. Making a 1:1000 dilution of the spike solution in a separate tube
 2. Adding 5 μ l of this 1:1000 dilution (10 pg/ μ l) to 1 μ g of total RNA sample
 3. Continuing with target preparation
- For 200 ng of mRNA, add 200 pg of spike dilution mix (at a concentration of 10 ng/ μ l). This represents a 1:1000 mass ratio of spike to polyA(+) RNA sample. Proceed by:
 1. Making a 1:100 dilution of the spike solution in a separate tube
 2. Adding 2 μ l of this 1:100 dilution (100 pg/ μ l) to 200 ng of mRNA sample
 3. Continuing with target preparation

Expression Report File from CodeLink Expression Analysis Software

The intensity data associated with each positive control probe is included in each Microsoft™ Excel output Expression Report file. For assistance with extracting and interpreting positive control probe performance, contact your Technical Service representative.

Distribution of Control Probes

For the best feedback on bioarray performance, controls must be evenly distributed across the bioarray; well-distributed controls remove location-

dependent bias and make performance assessments more sensitive. Figure 1 illustrates the dispense pattern of positive and negative controls on CodeLink™ Human Whole Genome bioarrays. Five unique probe sequences are represented on the bioarray for each of the six supplied bacterial transcripts. The bioarray also includes a panel of negative controls designed against *E. coli* genes, which were empirically tested to ensure no cross-reactivity in a large panel of different tissues from human, mouse, and rat. Each positive and negative control sequence is represented in a 12x redundancy on the bioarray surface, which improves the statistical power of these probe sets. The total number of positive control probes on the CodeLink™ Human Whole Genome Bioarray is 360, while the total number of negative control probes is 384.

Relative Signal Intensities

Accurate assessments of bioarray performance depend on control probes that demonstrate equivalent performance across and within all six bacterial transcripts. A large panel of positive controls was initially screened at two different spike levels: 1:256 000 and 1:64 000. Each probe-screening array contained six replicates of five unique probes per gene. Calculations were made across three bioarray replicates hybridized with the same cRNA sample. Figure 2 plots the average intensities for each probe, which are consistent across all transcripts. In addition, the tightness of the error bars (which represent one standard deviation of the mean of the five different probe hybridization intensities) further illustrates concordance of response for the different probe sequences.

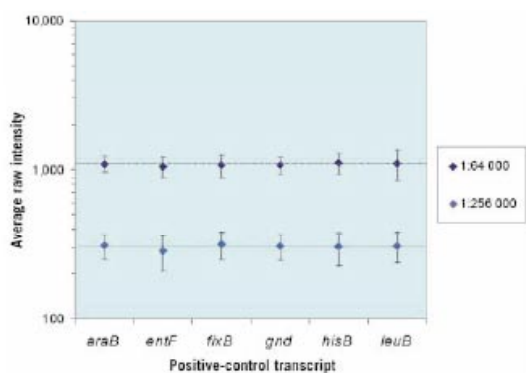


Fig. 2. Positive-control probes respond equivalently across the entire set of bacterial transcripts. We screened

positive-control probes containing all six bacterial cRNA spikes at mass ratios of 1:64 000 or 1:256 000 and plotted the average intensities for each bacterial transcript calculated from five unique probes. The horizontal lines represent the grand mean of each spike level. Error bars represent one standard deviation of the mean of the five different probe hybridization intensities.

Within-Slide Dynamic Range

To test how well the positive control probes measure a wide range of spike concentrations for each transcript, we spiked bacterial controls at the cRNA level across 3 logs of target concentration from 50 fM to 50 pM in four-fold increments corresponding to a spiked mass-ratio range of 1:2 048 000 to 1:2000 for the bacterial transcripts. Figure 3 displays the within-slide dynamic range for three different bioarray technical replicates that were independently spiked; a fourth, non-spiked bioarray was also plotted. In the absence of spike, the positive control probes yield signals below the negative control population (noise); when spiked, the probes offer a linear response across 3 logs of target concentration. Figure 3 also illustrates the high sensitivity of CodeLink bioarrays in accurately measuring transcripts in the low femtomolar range. Sensitivity at 50 fM (1:2 048 000 mass ratio) is well above noise, as illustrated by the standard error for *leuB* relative to the upper 99th percentile of the negative control population.

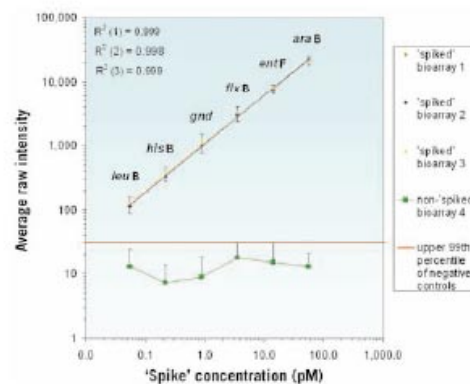


Fig. 3. Within-slide dynamic range derived from positive controls in CodeLink™ Human Whole Genome bioarrays. Bacterial transcripts for *leuB*, *hisB*, *gnd*, *fixB*, *entF*, and *araB* were 'spiked' into cRNA at 50 fM, 0.2 pM, 0.8 pM, 3.2 pM, 12.8 pM, and 51.2 pM, respectively. The average intensity was calculated across five unique probes for each bacterial transcript (probes are represented at a 12x redundancy within the bioarray, thus n = 60). Within-slide dynamic range is shown for three different bioarray technical replicates; a fourth bioarray was processed

without spiking and is illustrated by the horizontal green data points. Error bars represent one standard deviation of the average raw intensity for each bacterial transcript. The horizontal red line represents the upper 99th percentile of the negative controls within one of the bioarrays containing the spiked within-slide dynamic range.

Bioarray Signal Distribution and Relative Levels for Spiked Transcripts

The spiking procedure used in the within-slide dynamic range experiments (Fig. 3 and protocol) also illustrates the precision and localizations of the probe response. The probe response covers with precision the entire range of signal obtained using CodeLink™ Human Whole Genome bioarrays (Fig. 4). By following the spiking procedure outlined in the Protocol section of this document, CodeLink™ Human Whole Genome bioarray users should be able to duplicate the findings reported in Figures 3 and 4 for the array's positive controls, enabling them to directly assess sample quality and bioarray performance.

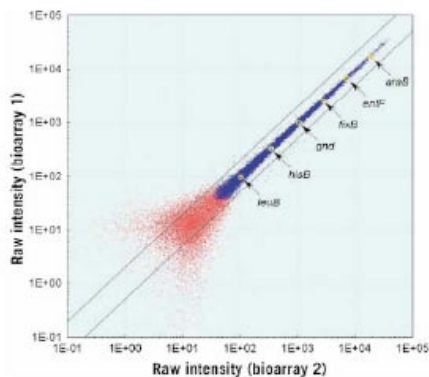


Fig. 4. Location and precision of positive-control spikes within the entire signal distribution for CodeLink™ Human Whole Genome bioarray probes. In the scatter plot above, all probe signals on the CodeLink™ Whole Genome Bioarray are represented between two technical replicates, hybridized with the same cRNA sample. Data points in red represent probes below noise on either bioarray replicate. The yellow circles illustrate the signal levels for each of the bacterial transcripts relative to the entire distribution of probe intensities, which indicates how the signal levels for the individual spiked transcripts performed relative to the entire population of probes (blue) and noise (red). The black 45° lines represent two-fold levels of differential expression.

References

1. CodeLink Gene Expression System: Manual Labeled cRNA Target Prep. Amersham Biosciences, 080074 (2004).
2. CodeLink Gene Expression System: Single Assay Bioarray Hybridization and Detection, Amersham Biosciences, 080075 (2004).
3. Ramakrishnan, R. et al. An assessment of Motorola CodeLink microarray performance for gene expression profiling applications, *Nucleic Acids Res.* 30, e30 (2002).

Applied Microarrays, Inc. ("AMI") reserves the right, subject to any regulatory approval if required, to make changes in specifications and features shown herein, or discontinue the product described at any time without notice or obligation. Contact your AMI Representative for the most current information. © 2007 Applied Microarrays, Inc. – All rights reserved. The AMI logo is a trademark of Applied Microarrays, Inc. Cy and CodeLink™ are trademarks of GE Healthcare companies. GenePix is a trademark of Axon Instruments Inc. QIAquick and RNEasy are trademarks of Qiagen. Agilent is a trademark of Agilent Technologies. Microsoft is a trademark of Microsoft Corp.

CodeLink Bioarrays are for research purposes only.