

The New CodeLink Whole Genome Bioarray platform: Enabling Technologies for high density, high performance oligonucleotide bioarrays

John McGowen, Hrisi Samartzidou*, Roxane Bonner, Randall Lockner, Richard Shippy, Tim Sendera, Gretchen Kiser, Tamma Kaysser-Kranich
Amersham Biosciences, 3200 W Germann Rd. Chandler, AZ 85248

Introduction

CodeLink™ Whole Genome Bioarray product offerings, including the newly launched CodeLink Human Whole Genome Bioarray, are the result of three main technological improvements: a new slide surface, production robotics for dispensing, and software for spot quantification and data analysis. These improvements enabled increased density and performance for developing a single 1 x 3 inch bioarray with the most comprehensive human genome coverage for gene expression profiling.

The purpose of this review will be to introduce the new slide surface that was co-developed with SurModics Inc. as a technology enabler for the CodeLink whole genome array platform. The new CodeLink high density (HD) substrate is a 3-D hydrophilic polymer matrix utilizing the same oligonucleotide attachment chemistry as CodeLink Activated Slides. The new slide surface exhibits the same high level of reproducibility, sensitivity, and specificity as with the original slides, but the new slide surface has been engineered to yield significant improvements in spot density that were necessary to achieve comprehensive coverage of the human genome.

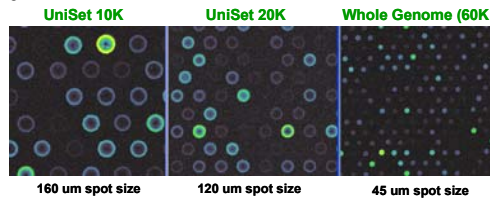


Figure 1: Representative post assay image showing decreasing spot size and improved morphology using the new CodeLink HD slides in the whole genome dispense format.

Materials and methods

Comparisons were made between the new CodeLink high density (HD) substrate and the original CodeLink Activated Slides (manufactured by SurModics Inc.) for both 20 K and whole genome (60 K) dispense formats using CodeLink UniSet™ Human 20K content.

All slides were assayed according to standard protocols using Cy™5-Streptavidin and were analyzed using CodeLink Expression Analysis v.4.0 software. All assay performance comparisons between slide types or platforms shown were processed in parallel on the same day with the same cRNA for a given experiment.

Reproducibility

Array-to-array reproducibility between slide types was assessed by calculating the coefficient of variation (Figure 2) and minimum detectable fold change (Table 1) for normalized signal intensities across three bioarrays for each slide type (10 µg cRNA, Human Placenta). The same high level of reproducibility is achieved with the 20K dispense format on CodeLink HD slides as is typical for the CodeLink UniSet 20K bioarray product offerings on the standard slide surface. In addition, when comparing normalized signal intensity array-to-array ratios across slide types, 98.7 % of probes were within two-fold.

Sensitivity

Known mass ratios of spiked bacterial positive controls were used to evaluate the dynamic response at the low end of the intensity spectrum, and to determine the detection limit of low copy number

transcripts (Figure 3). The CodeLink HD slides trend consistently higher in S/N relative to CodeLink Activated Slides.

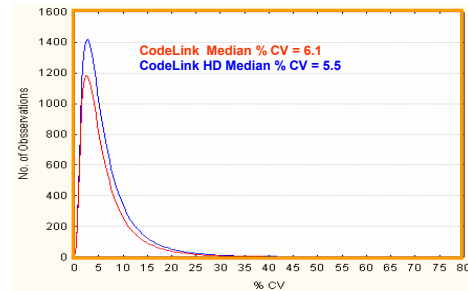


Figure 2: Log Normal fit showing comparison of %CV distribution between slide types with CodeLink UniSet Human 20K discovery probes. n = 3 bioarrays/slide type.

Table 1: MDFC Comparison for 20K dispense format (n = 3 bioarrays/slide type)

Description	# of Comp.	% Within 1.5 Fold	% Within 2 Fold	95 % Within x-fold
Total	43033	99.33	99.79	
Average	14344	99.33	99.8	1.24
+/- 95% Conf. Int.	126	0.0428	0.0084	

Description	# of Comp.	% Within 1.5 Fold	% Within 2 Fold	95 % Within x-fold
Total	53741	99.2	99.73	
Average	17914	99.2	99.73	1.24
+/- 95% Conf. Int.	13	0.0239	0.0144	

Across Slide Type Comparison				
Description	# of Comp.	% Within 1.5 Fold	% Within 2 Fold	95 % Within x-fold
Total	257850	94.37	98.73	
Average	17190	94.56	98.77	1.64
+/- 95% Conf. Int.	746	0.0096	0.0019	

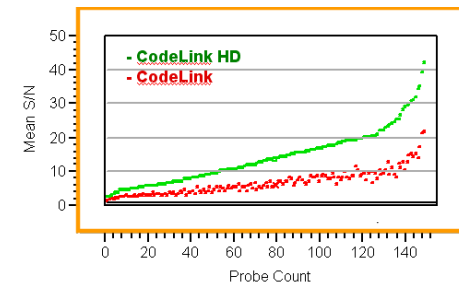


Figure 3: Comparison of Bacterial Control Probe mean signal/noise (signal = mean pixel intensity, noise = local background median +1.5 stdev). Graph represents a series of 150 bacterial probe sequences covering six transcripts; *araB*, *entF*, *fixB*, *gnd*, *hisB*, *leuB* (mRNA spiked into total RNA at an estimated mass ratio of 1:300 K).

Specificity

Three base mismatch control probes were used to assess specificity of the CodeLink HD slides and also to demonstrate that the availability of the entire probe sequence for hybridization is the same as on CodeLink Activated Slides (Figure 4).

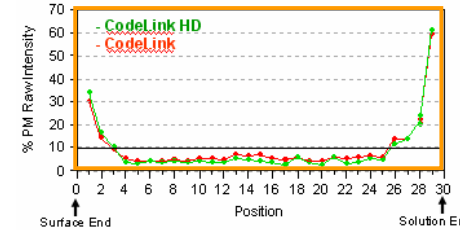


Figure 4: Specificity and the availability of the entire sequence to discriminate a 3-base mismatch is shown as a function of Mean % perfect match (PM) Raw Intensity vs. position of 3-base mismatch along the sequence (n = 18). Surface end indicates the 5' end of the probe.

Ratio Back Compatibility CodeLink vs. CodeLink HD

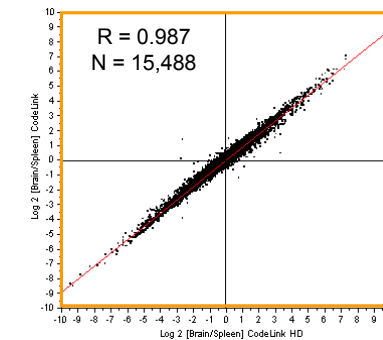


Figure 5: Assessment of differential expression ratio correlation between slide types printed at 20K density for Log₂ ratios of Brain over Spleen (N = 3 slides for each slide type per tissue). The correlation (Pearson's) is high (0.987) and the total number of ratios is shown.

Whole genome dispense format

Performance of the new CodeLink HD slides is comparable with the original slides when printed at 20K dispense density. Comparisons were also made using the new dispense robotics and the 60 K spot density of the CodeLink Whole Genome Bioarray platform. CodeLink Human Whole Genome Bioarrays were assayed along with CodeLink UniSet Human 20K Bioarrays, and a comparison of reproducibility (MDFC, Table 2) and correlation of differential expression ratios to the lower density platform were made (Figure 6) for the 20K subset of genes on the CodeLink Human Whole Genome Bioarrays. Sensitivity of the new platform was also assessed using bacterial control probes (representing six bacterial transcripts) in a cRNA spiking experiment, and sensitivity down to 1:2 048 000 mass ratio (0.05 pM) was demonstrated (Figure 7).

Table 2: MDFC UniSet Human 20K and 20K subset of HWG.

CodeLink: UniSet Human 20K (n = 3 slides)				
Description	# of Comp.	% Within 1.5 Fold	% Within 2 Fold	95 % Within x-fold
Total	41795	98.38	99.61	
Average	13932	98.38	99.61	1.33
+/- 95% Conf. Int.	62	0.0212	0.0158	

CodeLink HD: 20K Subset of HWG (n = 4 slides)				
Description	# of Comp.	% Within 1.5 Fold	% Within 2 Fold	95 % Within x-fold
Total	106061	98.23	99.38	
Average	17677	98.23	99.38	1.29
+/- 95% Conf. Int.	76	0.0037	0.0024	

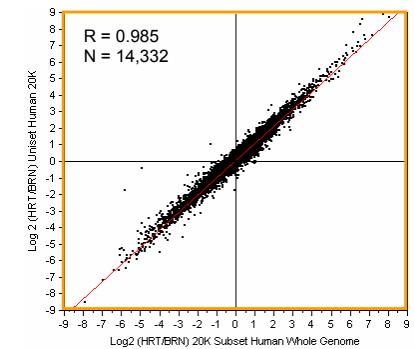


Figure 6: Assessment of differential expression ratio correlation between UniSet Human 20K bioarray and 20K subset of HWG Array for Log₂ ratios of Heart (HRT) over Brain (BRN). N = 3 slides for each product type per tissue.

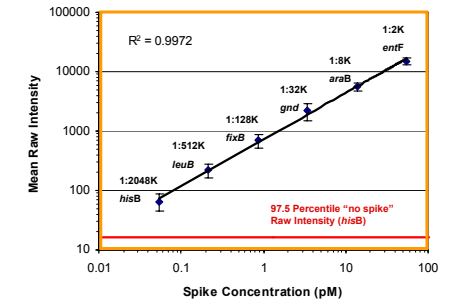


Figure 7: CodeLink Human Whole Genome Arrays show sensitivity of detection down to 1:2 048 000 mass ratio as determined by cRNA spiking experiments across three logs of concentration change (0.05 -50 pM). N = 180, error bars are +/- 1 stdev.

Conclusions

CodeLink bioarray platform advancements enable the deposition of probes covering the whole human genome on a single slide. This new high-density platform maintains the same performance as the current platform, including very high reproducibility, sensitivity, and dynamic range (0.05 - 50 pM). In addition, excellent correlation with the current platform for 20K bioarrays will allow users back compatibility of ratio data.

* To whom all correspondence should be addressed.
Hrisi.samartzidou@amersham.com

CodeLink, UniSet and Cy are trademarks of Amersham Biosciences Ltd Amersham and Amersham Biosciences are trademarks of Amersham plc. ©Amersham Biosciences Corp 2004- All rights reserved. All goods and services are sold subject to terms and conditions of sale of the company within the Amersham Biosciences group which supplies them. A copy of these terms and conditions are available on request. CodeLink Bioarrays are for research purposes only. Amersham Biosciences UK Limited, Amersham Place, Little Chalfont, Buckinghamshire England U.K. HP7 9NA. Amersham Biosciences AB SE-751 84 Uppsala Sweden. Amersham Biosciences Corp 800 Centennial Avenue PO Box 1327 Piscataway NJ 08855 USA. Amersham Biosciences Europe GmbH Munzinger Strasse 9 D-79111 Freiburg Germany