CodeLink Human Whole Genome Bioarray

55,000 human gene targets on a single bioarray

The CodeLink™ Human Whole Genome Bioarray comprises one of the most comprehensive coverages of the human genome, as it is known today, for genome-wide gene expression analysis on a single bioarray. CodeLink™ Human Whole Genome Bioarray targets ~57,000 transcripts and ESTs, including ~45,000 well characterized human gene and transcript targets based on the NCBI/UniGene database. This new bioarray is fully integrated with the CodeLink™ Expression Bioarray System of optimized reagents, protocols, processing tools, and software for ease of use and superior performance (Fig 1). Key features include:

• Accurate and reproducible biological results. ~55,000 quality probes were designed and experimentally tested. For best results with every bioarray experiment, our six-sigma manufacturing process and advanced non-contact deposition robotics ensure probe and deposition quality compliance.

• High assay sensitivity and specificity. An improved and proprietary 3-D aqueous gel slide surface enables efficient interaction between probe and target. With this new surface, sensitivity of detection is as low as 1:2,000,000 mass ratio determined by spiking experiments at the cRNA level (Table 1 and Fig 2) for enhanced detection of low abundance transcripts.

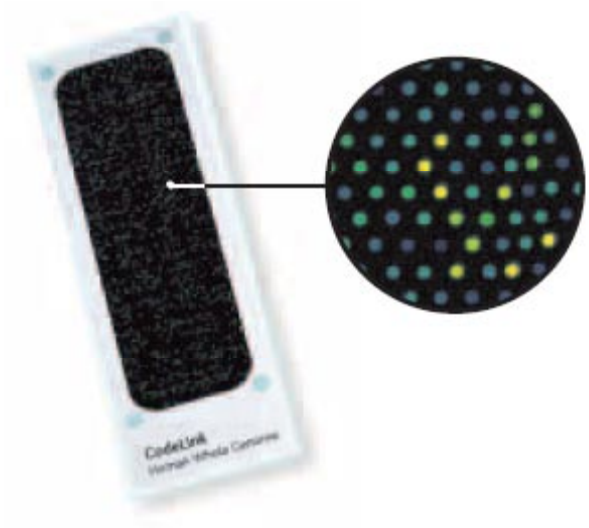

Fig 1. CodeLink™ Human Whole Genome Bioarrays are packaged with a unique Flex Chamber that allows simple setup, easy loading of target, and solution mixing for improved hybridization efficiencies. Optimized surface chemistry and non-contact deposition robotics result in excellent spot morphology and increased density.

CodeLink™ Human Whole Genome Bioarray performance

	Typical performance	*Benefit*
Sensitivity	Mass ratio < 1:1 000 000 with bacterial spiking into cRNA or total RNA level (Fig 2)	Detection of more low abundance transcripts (< 0.3 copies per cell)
Dynamic range	Linear signal response across three orders of magnitude (0.05 pM – 50 pM) (Fig 2)	More usable/accurate data across a broad range of gene expression levels
Specificity	Discrimination of > 2 base mismatches	Better discrimination of highly homologous genes
Signal reproducibility	Typically < 15% total median CV among production batches, all probes included; ~8% median CV among production batches, only probes above noise included	Reproducible differential expression results in every experiment (Fig 3 and Fig 4)
Minimum detectable fold change	> 98% within two-fold change for replicates	Reduced noise for better detection of small changes in gene expression and more usable data per bioarray (Fig 3 and Fig 4)

Table 1. CodeLink™ Human Whole Genome Bioarray observed performance specifications.

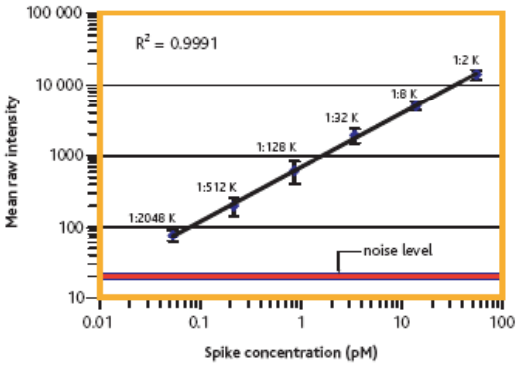

Fig 2. CodeLink™ Human Whole Genome Bioarrays show sensitivity of detection down to ~1:2,000,000 mass ratio, determined by spiking experiments at the cRNA level. Total dynamic range shown is 0.05-50 pM spike concentration range.

~57,000 Human transcripts represented in a single experiment

CodeLink™ Human Whole Genome Bioarrays target most of the known and predictive genes of the human genome as it is described today in the public domain. It is comprised of approximately 55,000 30-mer probes designed to conserved exons across the transcripts of targeted genes (Table 2). These 55,000 probes represent well annotated, full length, and partial human gene sequences from major public databases. CodeLink™ Human Whole Genome Bioarray probe sequences were selected from the NCBI UniGene build #165, RefSeq database (January 5, 2004 release) and dbEST database (January 8, 2004 release) (Table 3).

Well-annotated mRNA or coding sequences were chosen to ensure usefulness for a large range of applications in basic research, biotechnology, and

drug development (Table 4). Each sequence was carefully screened to ensure high-quality, specific probe design, and to reduce redundancy of gene targets. All of the probes designed to these sequences were functionally tested, and over 80% were functionally validated against 25 human tissues to ensure best representation of the gene and biologically relevant results.

CodeLink™ Human Whole Genome Bioarray specifications	
Number of arrays in set	1
Number of probes	54,841
Number of transcripts	57,347
Number of discovery genes	45,674*
Positive controls	360
Negative controls	384
Housekeeping genes	100
Number of probes per gene	One specific & functionally validated probe
Oligonucleotide probe length	30-mer probe
Probe proximity to 3' end	90% within 1000 bases; average distance 424 bases
RNA sample input	0.2–2 µg total RNA
Type of assay	One color
Storage & handling temperature	Room temperature

*Based on unique UniGene IDs.

Table 2. CodeLink™ Human Whole Genome Bioarray specifications.

CodeLink™ Human Whole Genome Bioarray probe distribution	
Unique NCBI accession numbers	57,841
Unique UniGene IDs	45,674
Unique RefSeq IDs	18,882
Unique locus link IDs	18,864
Ensembl genes	14,812

Table 3. CodeLink™ Human Whole Genome Bioarray content source and representation of public databases.

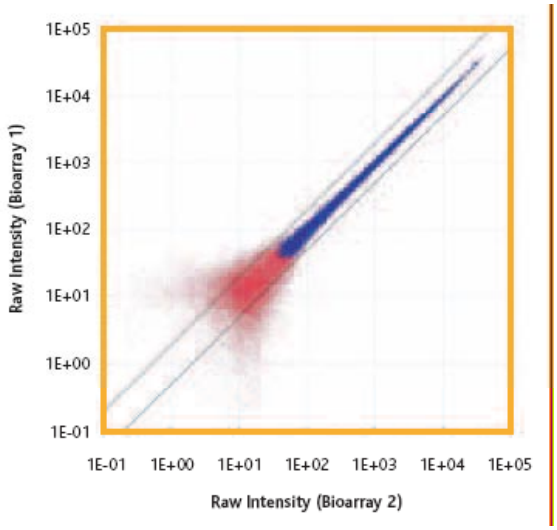

Fig 3. CodeLink™ Human Whole Genome Bioarrays exhibit excellent array-to-array signal intensity reproducibility.

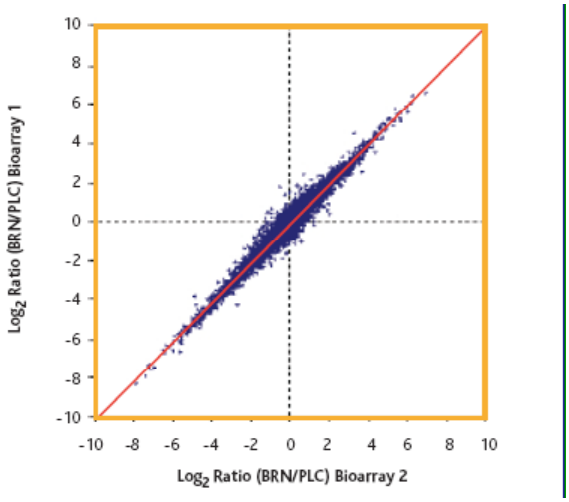

Fig 4. CodeLink™ Human Whole Genome Bioarrays ensure high differential expression ratio reproducibility for robust expression profiling. Ratios for brain vs placenta samples were calculated in this experiment.

As a result of their high-quality probe content and design, CodeLink™ Human Whole Genome Bioarrays are useful for performing differential expression studies of many important biological processes in a single experiment (Table 4). For instance, CodeLink™ Human Whole Genome Bioarray can be used to study unique gene expression profiles at different stages of oncogenesis. These studies help to characterize and understand tumor types and tumor

progression for use in drug target identification and personalized medicine. In addition, this bioarray enables the investigation of cell and tissue reactions to their environment from stimulus reception to signal transduction processing and response. This bioarray comprehensively covers cell differentiation, development, apoptosis, and metabolism for drugs and other cellular biomolecules. Expression profiling of the genes involved in these types of biological processes can be directly applied to target identification, pathway characterization, and drug effect prediction in basic research and drug development.

Back-compatible data

For flexibility and data preservation, CodeLink™ Human Whole Genome Bioarrays are compatible with lower density CodeLink™ human bioarrays. This means that existing data from CodeLink™ UniSet™ Human Bioarrays can be continuously used in future studies along with newly generated data from CodeLink™ Human Whole Genome Bioarrays (Fig 5).

CodeLink Expression Bioarray System

As part of the CodeLink™ Expression Bioarray System (Fig 6), CodeLink™ Human Whole Genome Bioarrays are compatible with existing instruments, reagents, software, and accessories.

• The CodeLink™ iExpress Assay Reagent Kit provides the required reagents and easy-to-follow protocol for target preparation, including cDNA synthesis and *in vitro* transcription. This high-yield reagent kit produces sufficient amplified material for multiple replicate experiments.

• The new CodeLink™ Gene Expression Analysis v5.0 software supports all current CodeLink™ bioarrays, including CodeLink™ Whole Genome Bioarrays. This latest software version continues to provide accurate data extraction and integration with manufacturing quality control information and gene information with spot-finding and primary data analysis. In addition, v5.0 introduces improved data reports and the ability to normalize using alternate methods, including housekeeping genes and custom probe sets. This is especially important for smaller focused array sets, such as the iExpress 16-assay focused arrays. Data generated with CodeLink™ Gene Expression Analysis v5.0 software is back-compatible with data generated with previous software versions.

• A complete line of accessories, including parallel processing and hybridization tools, is available for use with the CodeLink™ Human Whole Genome Bioarray.

CodeLink Human Whole Genome gene classification

Major biological process	No. of genes present
Oncogenesis	~1,200
Cell cycle	~1,400
Cell-cell signaling	~1,000
Signal transduction	~3,000
Metabolism	~2,000
Developmental processes	~1,400
Transcription and translation	~2,700
Immune and inflammation response	~1,100
Protein phosphorylation	~800
Apoptosis	~600
RNA processing	~550
Ion transport	~1,150
Protein transport	~450
Electron transport	~500
Synaptic transmission	~400
Kinases	~200

Table 4. The biological processes listed are approximate and conservative estimates of the minimum number of genes targeted in each CodeLink™ Human Whole Genome Bioarray for study. The above classification is based on the "BIOLOGICAL PROCESS" organizing principle from Gene Ontology™ Consortium.

-

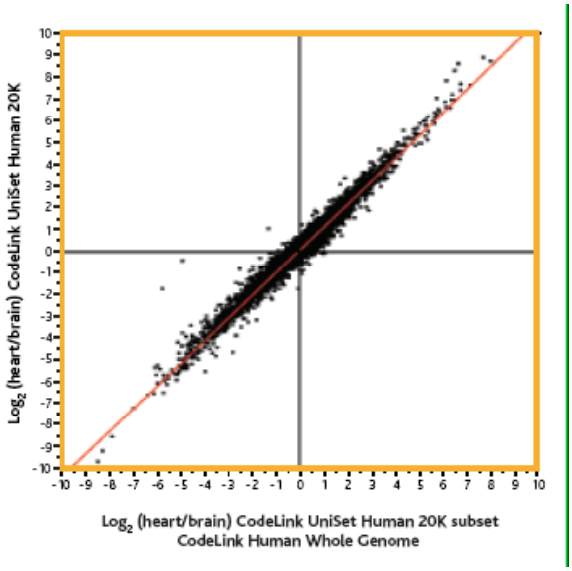

Fig 5. CodeLink™ Human Whole Genome Bioarrays generate data that is compatible with corresponding data from lower density bioarrays, including CodeLink™ UniSet Human 20K I bioarrays. The differential expression ratios of brain over placenta show high correlation between the two CodeLink™ bioarray platforms

Fig 6. The CodeLink™ Human Whole Genome Bioarray is fully integrated with the CodeLink™ Bioarray System.

Ordering information

Description	Size	Code number
CodeLink Human Whole Genome Bioarray	6/pack	300026-6PK
CodeLink iExpress Assay Reagent Kit	24 reactions	67601000
Universal Shaker Kit	12 slides	310031
Expression Parallel Processing Kit	12 slides	310010
Cy™5-Streptavidin for Microarrays	1 mg	28900224
CodeLink™ Gene Expression Analysis v5.0 software		310035

For further information, please contact the Applied Microarrays, Inc. sales office below.

online www.appliedmicroarrays.com

Applied Microarrays, Inc.
7700 S. River Parkway
Tempe, AZ 85284 USA
Email: sales@appliedmicroarrays.com
Phone : (480) 775 6320 FAX: (800) 927 9315
www.appliedmicroarrays.com