AFFYMETRIX®

Technical Note

Design and Performance of the GeneChip® Human Genome U133 Plus 2.0 and Human Genome U133A 2.0 Arrays

The GeneChip® Human Genome U133 Plus 2.0 (HG-U133 Plus 2.0) and GeneChip Human Genome U133A 2.0 (HG-U133A 2.0) Arrays represent the latest evolution in the Affymetrix® GeneChip Expression Analysis Platform, offering the most comprehensive tool for the analysis of genome-wide expression on a single array.

The HG-U133 Plus 2.0 Array includes all of the probe sets from the previous generation HG-U133A and HG-U133B Arrays, as well as nearly 10,000 additional probe sets. The increased feature density of the HG-U133 Plus 2.0 Array offers researchers the ability to measure transcription over the entire human genome in a single hybridization. Greater than 54,000 probe sets are used to analyze the expression level of more than 47,000 transcripts and variants, including approximately 38,500 well-characterized human genes.

The HG-U133A 2.0 Array also leverages this high-density design to include the content of the previous HG-U133A Array on a smaller format for more efficient analysis. More than 22,000 probe sets are employed to analyze the expression level of over 18,400 transcripts and variants, including more than 14,500 well-characterized human genes.

Extensive testing on both arrays revealed that data from the HG-U133A Array are similar to data generated on the higher density HG-U133 Plus 2.0 and HG-U133A 2.0 Arrays. In terms of assay sensitivity, control spikes were routinely detected in a complex background at a transcript concentration of one copy per 200,000.

Introduction

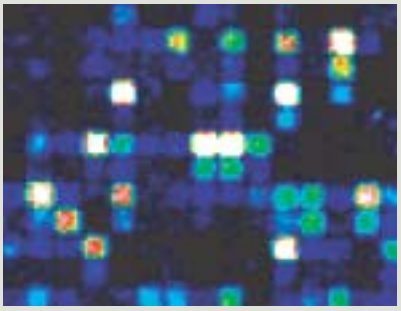
This Technical Note summarizes the design and data from a number of studies that were designed to evaluate the new GeneChip® Human Genome U133 Plus 2.0 (HG-U133 Plus 2.0) and GeneChip Human Genome U133A 2.0 (HG-U133A 2.0) Arrays. Advances in array technology enable synthesis of these arrays with 11-µm features for higher density content than the previous 18-µm arrays. An improved manufacturing process provides sharper features, even with the decrease in overall feature size (see Figure 1). Used in conjunction with the new GeneChip Scanner 3000, enabled for high-resolution scanning at 1.56 µm pixel resolution, scientists can now bring the latest advances in array technology to their research.

Comparisons to the previous HG-U133A and HG-U133B designs were straightforward in that the same probe sets, array design strategy, and assay are utilized for the new HG-U133 2.0 Arrays. Like all GeneChip brand arrays, the new HG-U133 Plus 2.0 and HG-U133A 2.0 Arrays utilize probe sets to provide multiple, independent measurements for each transcript. The use of probe sets provides the most stable platform for accurate, reliable, and reproducible results.

The data presented here, as well as numerous developmental studies, demonstrate highly comparable results generated with both the current and previous generation HG-U133 expression arrays. The HG-U133 Set was scanned with the GeneArray® Scanner, and data analysis was performed with Affymetrix® Micorarray Suite software (MAS) 5.0 using the default Statistical Algorithm parameters. Data acquisition for the HG-U133 Plus 2.0 and the HG-U133A 2.0 Arrays was performed using the GeneChip Scanner 3000, enabled for high-resolution scanning. Data analysis for the HG-U133 Plus 2.0 and HG-U133A 2.0 Arrays was per-

Figure 1. Improved feature definition. **A.** Image from a portion of a GeneChip® Human Genome U133 Plus 2.0 .dat file, showing 11-µm features scanned with the GeneChip Scanner 3000. **B.** Image from a portion of a Human Genome U133A .dat file, showing 18-µm features scanned with the GeneArray® 2500 Scanner. Note the improvement in the sharpness, definition, and signal uniformity of the features in Panel A.

Panel A. HG-U133 Plus 2.0 Array



Panel B. HG-U133A Array

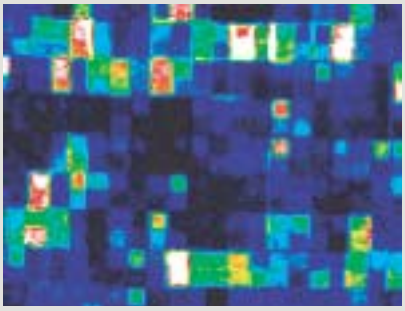


Table 1. Sources and numbers of sequences used for the GeneChip® Human Genome U133 Plus 2.0 Array design. UniGene clusters were used as a starting point for the design process but were not used as the main source of sequence information. The use of primary sequence sources provided better control over the regions used and access to additional annotation information, such as sequence quality parameters from dbEST. For the oringinal HG-U133 Set content, a draft assembly of the human genome from the University of California, Santa Cruz (April 2001) was used to improve cDNA sequence orientation and annotation. For the additional HG-U133 Plus 2.0 Array content, a draft assembly of the human genome from NCBI (Nov. 2002, #31) was used for the same purpose.

	Release Date		Clusters or Sequences		Used in Design		
Source	**HG-U133**	**HG-U133 Plus 2.0**	**HG-U133**	**HG-U133 Plus 2.0**	**HG-U133**	**HG-U133 Plus 2.0**	**Total**
UniGene	April 2001 (#133)	Jan. 2003 (#159)	95,569	108,944	31,729	7,933	38,572*
dbEST	April 2001	Feb. 2003	3,471,886	5,030,353	2,619,747	49,449	2,669,196
GenBank	April 2001 (#123.0)	Feb. 2003 (#134.0)	61,523	112,688	38,168	10,967	49,135
RefSeq	April 2001	March 2003	12,716	18,599	12,461	1,231	13,692
Total Sequences			**3,546,125**	**5,161,640**	**2,670,376**	**61,647**	**2,732,023**

*In a small number of cases, the same UniGene identifier was present in both sets and was only counted once in the total.

formed with GeneChip® Operating Software (GCOS), the High-Resolution Scanning Patch, and default Statistical Algorithm parameters.

Some general performance parameters, such as Signal correlation, demonstrate a high degree of similarity between the new HG-U133 2.0 Arrays and the previous HG-U133 Arrays. In terms of the percentage of Present calls observed in tissue panel studies, the HG-U133 Plus 2.0 and HG-U133A 2.0 Arrays produced higher call rates than the HG-U133 Arrays. The differences in Detection call rates were less pronounced among the probe sets found in the HG-U133A Array than the HG-U133B Array.

Sensitivity experiments, utilizing spikes known to be absent in the pool of background RNA, also defined the performance of the new arrays. Using the HG-U133 Plus 2.0 and the HG-U133A 2.0 Arrays, spikes were readily detected at a concentration of 0.75 pM, a level calculated to be approximately one transcript per 200,000. At 0.75 pM levels, two-fold increases in concentration were also routinely detected using the Comparison Algorithm within GCOS software.

GeneChip® Human Genome U133 Plus 2.0 Array Design

The new HG-U133 Plus 2.0 Array leverages the same array design strategy as the previous HG-U133 Set. All of the probe sets from the initial HG-U133A and B Arrays are used in the new HG-U133 Plus 2.0 Array and are combined with an additional 9,921 probe sets. The result is the most comprehensive coverage of the known human transcriptome on a single microarray. The HG-U133A 2.0 Array contains all of the probe sets from the HG-U133A Array.

Various public data sources were used for the HG-U133 Plus 2.0 Array design (Table 1). Sequence data for new content were obtained from dbEST (NCBI, Feb. 2003), GenBank® (NCBI, Feb. 2003, Release 134), and RefSeq (NCBI, March 2003). Additionally, a draft assembly of the human genome (NCBI, Nov. 2002, Build 31) was used to assess sequence orientation and quality. The initial sequence curation process involved:

- Collection of sequences and annotations from various public sources
- Identification and removal of vector sequences
- Sequence alignment to the draft assembly of the human genome
- Detection of polyadenylation sites
- Orientation of sequences, using consensus splice sites from genome alignments, detected polyadenylation sites, coding sequence (CDS), and EST read direction annotations
- Identification and removal of low-quality regions of EST sequences

UniGene (NCBI, Jan. 2003, Build 159) was then used to create initial clusters of cDNA sequences (Figure 2). Sequence-based subclustering was accomplished using StackPACK software (Electric

Figure 2. Process of clustering cDNA sequences. Sequence cluster information from UniGene was used to create initial seed clusters. Seed clusters were subclassified into one or more StackPACK subclusters with assemblies using StackPACK (Electric Genetics). Subclusters with orientation problems were further subclustered into orientation subclusters, which were then processed by StackPACK.

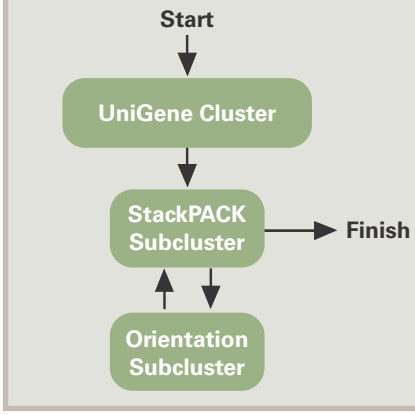
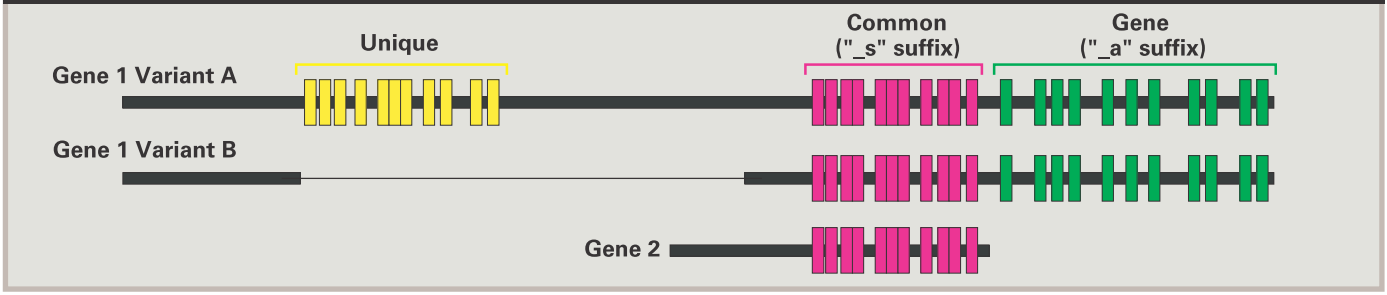


Figure 3. Different probe set types are indicated by suffices to the probe set name. Unique probe sets are predicted to perfectly match only a single transcript. Gene probe sets, with an "_a" suffix, are predicted to only perfectly match transcripts from the same gene. Common probe sets, with a "_s" suffix, are predicted to perfectly match multiple transcripts, which may be from different genes. Probe sets that have a "_x" suffix are not shown here but are described in the text.



Genetics). This step partitions alternative transcript isoforms into separate clusters and helps to remove problematic sequences. In some cases, sequence-based subclusters were further subgrouped due to conflicting orientation calls within the subcluster assembly. To be conservative when selecting probes, at least 75 percent identity in all of the member sequences was required when creating a consensus sequence.

Probes were selected from the 600 bases most proximal to the 3' end of each transcript, and probe selection regions were defined using any of the following criteria:

- 3' ends of RefSeq and complete CDS mRNA sequences (full length end)
- Eight or more 3' EST reads terminating at the same position (strong evidence for polyadenylation)
- 3' end of the assembly (consensus end)

This approach identified alternative polyadenylation sites internal to the assembly end. In contrast to the previous generation HG-U133 Array design, the probe selection region for the novel content on the HG-U133 Plus 2.0 Array was always selected from the consensus sequence in order to simplify data analysis. When alternative polyadenylation sites were less than 600 bases apart, only the probe selection region based on the upstream polyadenylation site was used.

ADDITIONAL CONTENT ON THE HUMAN GENOME U133 PLUS 2.0 ARRAY

The new HG-U133 Plus 2.0 Array utilizes all of the probe sets from the HG-U133 Set, as well as additional probe sets designed by selecting probes in novel regions that did not match an existing HG-U133 probe set. Probes and probe sets were selected for the new probe selection regions as described by the *Array Design for the GeneChip® Human Genome U133 Set* Technical Note. In short, multiple linear regression models, with a strong thermodynamic component were used to predict probe performance. Eleven-probe-pair probe sets were then selected based on predicted probe characteristics, such as performance, uniqueness metrics, and spacing rules.

A new non-unique probe set type, "_a", was added to indicate those probe sets that recognize multiple alternative transcripts from the same gene (Figure 3). Probe sets with common probes among multiple transcripts from separate genes are annotated with a "_s" suffix. Note that the "_a" suffix was not used at the time the HG-U133A and HG-U133B Arrays were designed. For consistency, the names of existing probe sets with "_s" suffices were not changed. New probe sets designed from additional content include both "_s" and "_a" suffices. This is described in greater detail in the Technical Note entitled *Array Design and Performance of the GeneChip® Mouse Expression Set 430.*

Occasionally, it was not possible to select a unique probe set or a probe set with identical probes among multiple transcripts. In this case, similarity criteria were not applied and the resulting probe set is annotated with a "_x" suffix. Such probe sets will contain some probes that are identical or highly similar to other sequences. The probe set as a whole may cross-hybridize with other transcripts but should hybridize uniformly across probe pairs to the intended target. Data generated from "_x" probe sets should be interpreted with caution due to the likelihood that some of the Signal measurements for a subset of the probes in the probe set are from transcripts other than the one being intentionally measured.

The 9,921 new probe sets were combined with the previous generation HG-U133 probe sets to create the HG-U133 Plus 2.0 Array design (Table 2). The result is a single array with more than 54,000 probe sets representing approximately 38,500 genes (estimated by UniGene coverage).

Based on the information available when the arrays were designed, 70 percent of the probe sets represent subcluster assemblies containing one or more non-EST sequences. Of the 16,737 EST-based probe sets (from the original HG-U133 Array design, Table 2), approximately 9,000 probe sets can now be associated with an mRNA or other non-EST sequence. This updated annotation information is available through the NetAffx™ Analysis Center (www.affymetrix.com).

Table 2. Classification and number of probe sets on the GeneChip® Human Genome U133 Plus 2.0 Array design. It is estimated that this set interrogates over 47,000 transcripts from approximately 38,500 genes. The first tier provides a summary of content with regard to the listed metrics. The second tier provides a summary of probe set content based on annotation quality. The probe sets are assigned to the classifications based on the sequence quality of the subcluster (Full Lengths, Non-ESTs, ESTs) and the justification for the region from which probes were selected (based on strong evidence for polyadenylation, full length end, and/or consensus end). Classification was based on information available at the time of the array design. Probe set classifications for the HG-U133A and HG-U133B Array designs are shown for comparison.

Classification	HG-U133 Plus 2.0	HG-U133A 2.0 HG-U133A	HG-U133B
Probe Sets	54,120	21,722	22,577
UniGene Clusters	38,572	14,564	19,298
Additional Potential Full Lengths	841	517	211
Subclusters	47,401	18,445	21,070
Full Lengths			
Full Length End and Strong Evidence for Polyadenylation	5,823	5,360	447
Strong Evidence for Polyadenylation	2,240	1,624	603
Full Length End	12,064	8,415	1,219
Consensus End	509	314	143
Non-ESTs (excluding Full Lengths)			
Strong Evidence for Polyadenylation	3,300	1,477	1,657
Consensus End	13,459	3,057	3,246
ESTs			
Strong Evidence for Polyadenylation	6,285	1,054	5,242
Consensus End	10,440	421	10,020

Performance of the GeneChip® Human Genome U133 Plus 2.0 and Human Genome U133A 2.0 Arrays

TISSUE PANEL STUDIES

A tissue panel survey was used to compare performance between the HG-U133A Array and the new HG-U133 Plus 2.0 and HG-U133A 2.0 Arrays. Five human tissue samples were hybridized in triplicate to HG-U133A, HG-U133A 2.0, and HG-U133 Plus 2.0 Arrays. Metrics included overall percentage of probe sets called Present, Detection call concordance, and Signal correlation.

Table 3 summarizes the percentage of probe sets called Present for the five tissues on the HG-U133A, HG-U133B, HG-U133A 2.0, and HG-U133 Plus 2.0 Arrays. In general, the HG-U133A 2.0 and the HG-U133 Plus 2.0 Arrays produced a greater percentage of Present calls than the HG-U133A Array for those probe sets in

Table 3. Percentage of Present calls. Five human tissue samples were hybridized in triplicate to GeneChip® HG-U133A, HG-U133B, HG-U133A 2.0, and HG-U133 Plus 2.0 Arrays. The average percentage of probe sets called Present is shown for each array type. The HG-U133 Plus 2.0 Array is represented with three columns. The left column lists the percentage of Present calls for the probe sets shared by the HG-U133 Plus 2.0 Array and HG-U133A Array (21,722 probe sets). The next column lists the percentage of Present calls for the probe sets shared by the HG-U133 Plus 2.0 Array and the HG-U133B Array (22,577 probe sets). The right column shows the percentage of Present calls for all 54,120 probe sets on the HG-U133 Plus 2.0 Array.

	Percent Present Calls					
	HG-U133A	HG-U133B	HG-U133A 2.0	HG-U133 Plus 2.0		
Tissue				HG-U133A Probe Sets	HG-U133B Probe Sets	All Probe Sets
Brain	48.9%	30.4%	55.2%	55.2%	39.4%	40.5%
Heart	50.8%	26.9%	53.8%	52.3%	32.8%	35.8%
Kidney	52.1%	28.2%	54.8%	55.2%	35.3%	38.1%
Liver	45.0%	20.1%	48.9%	48.2%	27.8%	31.9%
Prostate	60.2%	30.8%	59.5%	56.3%	36.3%	39.1%

common. The lone exception on Table 3 is seen in prostate tissue. In addition, the percentage of Present calls for the two new arrays exhibited greater similarity to each other than to the HG-U133A Array.

The same trend was observed when common probe sets between the HG-U133 Plus 2.0 Arrays and the HG-U133B Array were compared; however, in this case, the probe sets on the new arrays clearly produced a higher percentage of Present calls (5.9 to 9.0 percent greater Present calls). The percentage of Present calls for all of the probe sets on the HG-U133 Plus 2.0 Array are also shown to provide users with an expected range when first using the array.

Detection call concordance between the HG-U133A Array and the HG-U133 Plus 2.0 and HG-U133A 2.0 Arrays was also assessed. This metric is used to evaluate the agreement in Detection calls when identical samples are hybridized to different array designs. Only the probe sets in common to the pairs of arrays being compared were used for this analysis. Table 4 summarizes the results for five human tissue samples.

Detection call concordance ranged from 90 to 93 percent. The Detection call concordance was higher between the HG-U133A 2.0 and HG-U133 Plus 2.0 Arrays, which share the same feature size and manufacturing methods (92 to 93 percent). Comparisons between the HG-U133A Array and the new HG-U133 Plus 2.0 and HG-U133A 2.0 Arrays gave slightly lower concordance values. The majority of the discordant probe sets were classified as Present on the HG-U133 Plus 2.0 and the HG-U133A 2.0 Arrays and classified as Absent on the HG-U133A Array. These discordant calls are explained by the fact that the HG-U133 Plus 2.0 and the HG-U133A 2.0 Arrays had a higher percentage of Present calls than the HG-U133A Array (Table 3). Overall, there is good agreement in the Detection calls, with a slightly higher percentage of Present calls on the HG-U133A 2.0 and the HG-U133 Plus 2.0 Arrays.

Table 4. Detection call concordance. Five human tissue samples were hybridized in triplicate to the GeneChip® HG-U133A Array, as well as to the HG-U133A 2.0 and the HG-U133 Plus 2.0 Arrays. Detection call concordance was calculated by comparing, in a pair-wise manner, each of the replicates hybridized to the first array listed to each of the replicates hybridized to the second array in the first column. Percent concordance was defined as the average percentage of probe sets giving concordant Present or concordant Absent calls for the three pair-wise comparisons between array types. For this analysis, Marginal calls were defined as Absent.

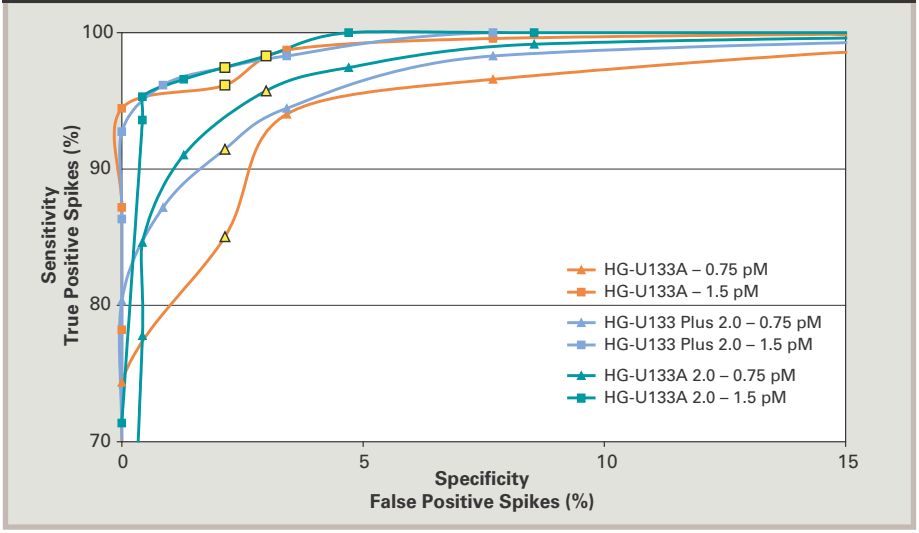
Brain	**Percent Concordance**
HG-U133A vs. HG-U133A 2.0	90%
HG-U133A vs. HG-U133 Plus 2.0	90%
HG-U133A 2.0 vs. HG-U133 Plus 2.0	93%
Heart	**Percent Concordance**
HG-U133A vs. HG-U133A 2.0	91%
HG-U133A vs. HG-U133 Plus 2.0	91%
HG-U133A 2.0 vs. HG-U133 Plus 2.0	93%
Kidney	**Percent Concordance**
HG-U133A vs. HG-U133A 2.0	92%
HG-U133A vs. HG-U133 Plus 2.0	92%
HG-U133A 2.0 vs. HG-U133 Plus 2.0	93%
Liver	**Percent Concordance**
HG-U133A vs. HG-U133A 2.0	90%
HG-U133A vs. HG-U133 Plus 2.0	90%
HG-U133A 2.0 vs. HG-U133 Plus 2.0	92%
Prostate	**Percent Concordance**
HG-U133A vs. HG-U133A 2.0	93%
HG-U133A vs. HG-U133 Plus 2.0	92%
HG-U133A 2.0 vs. HG-U133 Plus 2.0	93%

Table 5. Signal correlation. Five human tissues were run in triplicate on the GeneChip® HG-U133A, HG-U133A 2.0, and HG-U133 Plus 2.0 Arrays. The Signal values of the probe sets in common between pairs of arrays were used to calculate r^2 values. The three values obtained from comparing the pairs of replicates were averaged to give the values in the table. The first column of data shows average r^2 values for comparisons between the HG-U133A and the HG-U133A 2.0 Arrays. The second column of data shows average r^2 values for comparisons between the HG-U133A and the HG-U133 Plus 2.0 Arrays. The third column shows average r^2 values for comparisons between the HG-U133A 2.0 and the HG-U133 Plus 2.0 Arrays.

Tissue	**HG-U133A vs. HG-U133A 2.0 r^2**	**HG-U133A vs. HG-U133 Plus 2.0 r^2**	**HG-U133A 2.0 vs. HG-U133 Plus 2.0 r^2**
Brain	0.963	0.949	0.994
Heart	0.975	0.970	0.995
Kidney	0.972	0.961	0.995
Liver	0.975	0.967	0.995
Prostate	0.974	0.977	0.995

Signal correlation between the array types was evaluated over the tissue panel set. The data are presented in Table 5. For the five tissues sampled, r^2 values ranged from 0.95 to 0.98 when comparing the 18-µm HG-U133A Array to the new 11-µm HG-U133 Plus 2.0 and the HG-U133A 2.0 Arrays. These data demonstrate a high degree of Signal correlation between the same samples hybridized to the different arrays. The Signal correlation between the HG-U133 Plus 2.0 and the HG-U133A 2.0 Arrays was extremely high (r^2>0.99), again showing the high degree of similarity between the 11-µm designs.

Figure 4. Detection call sensitivity and specificity. This Receiver Operating Characteristic (ROC) curve plots Detection call results for 26 human spikes at two concentrations (0.75 pM and 1.5 pM) for each array type. Arrays include the GeneChip® HG-U133 Plus 2.0, HG-U133A 2.0, and HG-U133A Arrays. Three arrays from three manufacturing lots (nine hybridizations) were used for each spike pool concentration. The y axis represents the percentage of True Positive Spikes (Sensitivity), and the x axis represents the percentage of False Positive Spikes (Specificity), measured when transcripts are absent from the sample. Each of the six curves was generated by varying the Detection call *p*-value cutoff, Alpha1, and plotting percentages of true and false Detection calls for the two shown concentrations. Each point represents a different value of Alpha1. The yellow squares and triangles represent the default Detection call *p*-value cutoff, 0.05. The inverse relationship between true positive and false positive Detection is traced by each curve, with ideal performance corresponding to the upper left corner of the graph (highest true positive rate with lowest false positive rate).



SENSITIVITY AND REPRODUCIBILITY

To evaluate the performance of the HG-U133 Plus 2.0 and HG-U133A 2.0 Arrays relative to the HG-U133A design, a set of 26 human cRNA spikes was utilized. These spikes were shown to be absent by real-time RT-PCR analysis in the background RNA pool used for the studies. The spikes were evaluated at nine different concentrations in a modified Latin Square design. (Latin Square designs are discussed in greater detail in the *New Statistical Algorithms for Monitoring Gene Expression on GeneChip® Probe Arrays* Technical Note.) Each type of array was evaluated using three manufacturing lots.

To evaluate the Detection call sensitivity of the HG-U133A Array and the new HG-U133 Plus 2.0 and the HG-U133A 2.0 Arrays, the spike data from the Latin Square design were plotted as a Receiver Operating Characteristic (ROC) curve (Figure 4), which graphs the reciprocal relationship between sensitivity and specificity. In this ROC curve the True Positive rate (y axis) was plotted against the False Positive rate (x axis). Data from the three array designs are shown for 0.75 and 1.5 pM spike concentrations.

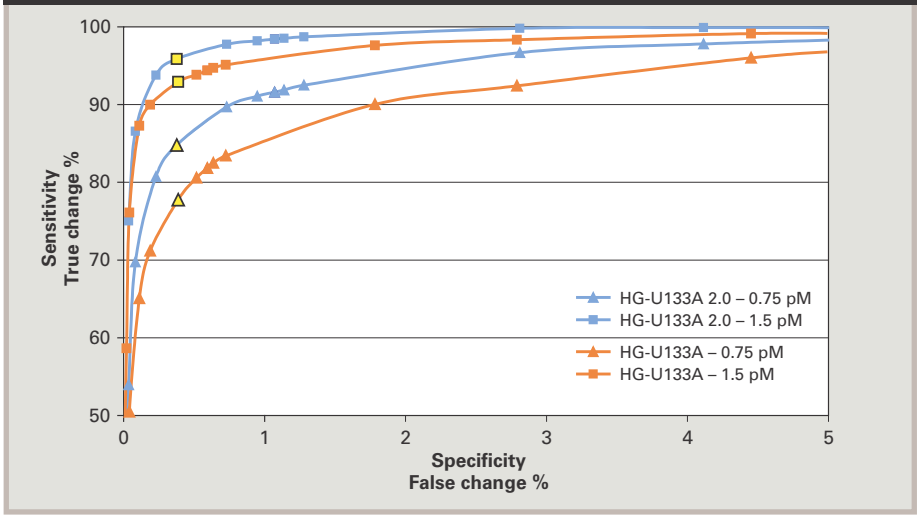
The curves for each of the three arrays at the spike concentration of 1.5 pM overlap, implying equivalent Detection call sensitivity for the three arrays. In the case of the 0.75 pM spikes, the curves produced by the HG-U133 Plus 2.0 and the HG-U133A 2.0 Arrays were resolved from the HG-U133A Array data. At this concentration, the HG-U133 Plus 2.0 and the HG-U133A 2.0 Arrays had a greater Detection call sensitivity in that more true positives were detected than for the HG-U133A Array with similar false positive rates.

Utilizing the default Alpha1 value for the Detection call algorithm, 1.5 pM spikes were detected with 97, 98, and 96 percent accuracy for the HG-U133 Plus 2.0, the HG-U133A 2.0, and the HG-U133A Arrays, respectively (yellow squares in Figure 4). The Detection call accuracy was 91, 96, and 84 percent for the HG-U133 Plus 2.0, the HG-U133A 2.0, and the HG-U133A Arrays, respectively, at the 0.75 pM spike concentration (yellow triangles in Figure 4).

The Change call sensitivity of the HG-U133A 2.0 and HG-U133A Arrays was also evaluated by ROC curves (Figure 5). In this case, the accuracy of detecting a two-fold higher spike concentration as an Increase call (y axis) was plotted against the False Change rate for the non-spike probe sets (x axis). Two comparisons were plotted for this study: spike concentrations of 1.5 pM were compared to spikes at 0.75 pM, and spikes at 3.0 pM were compared to spikes at 1.5 pM.

Figure 5 clearly illustrates the inverse relationship between the True Change call rate and the False Change rate; higher True Change call rates were obtained by increasing the tolerance for False Changes. The graph shows that for both pairs of comparison analyses, the new HG-U133A 2.0 Array curves approach the y axis maxi-

Figure 5. Change call sensitivity and specificity. This Receiver Operating Characteristic (ROC) curve plots Change call results for two sets of comparison analyses. The first comparison was between 26 human spikes at 1.5 pM compared to 0.75 pM; the second was for 26 human spikes at 3.0 pM compared to 1.5 pM. Arrays include the GeneChip® HG-U133A 2.0 (blue) and the HG-U133A (orange) Arrays. The y axis represents the percentage of True Change calls identified as Increase calls for spike probe sets only (Sensitivity); the x axis represents the percentage of False Change calls, measured for all probe sets except spike and control (AFFX prefix) probe sets in comparisons of arrays hybridized with identical sample (Specificity). False Change is defined as an Increase with a Signal Log Ratio ≥ 1 or a Decrease with a Signal Log Ratio ≤ -1. Each of the four curves is generated by varying the Change call *p*-value cutoffs, Gamma1L and Gamma1H, and plotting percentages of True and False Change calls for the two comparisons. The inverse relationship between Sensitivity and False Change is traced by each curve, with an optimal position in the upper left corner of the graph (highest True Change with lowest False Change). Each point represents the True and False Change for a different value of Gamma1L/H. The yellow symbols (squares and triangles) represent the performance resulting from the default values of Gamma1L/H for each array type (HG-U133A: 0.0045; HG-U133A 2.0: 0.002). As in the previous figure, three manufacturing lots with each sample analyzed in triplicate, were used to generate this graph.



mum at a faster rate than the HG-U133A Array. The HG-U133A 2.0 Array had a Change call sensitivity of 96 percent and 85 percent for 3.0 pM vs. 1.5 pM and 1.5 pM vs. 0.75 pM, respectively, using the default parameters (yellow symbols on blue lines in Figure 5). In contrast, the HG-U133A Array had a Change call sensitivity of 93 percent and 78 percent for 3.0 pM vs. 1.5 pM and 1.5 pM vs. 0.75 pM, respectively, again using default parameters (yellow symbols on orange lines in Figure 5). Both arrays had False Change rates of approximately 0.4 percent.

LOT-TO-LOT VARIABILITY STUDY

To assess the uniformity of assay performance with the HG-U133 Plus 2.0 Arrays, multiple manufacturing lots were tested with samples containing the same 26 control spikes derived from human cDNA clones. Sensitivity measurements based on the spikes and reproducibility measurements based on the non-spike, non-control (~54,000) probe sets were obtained for the nine lots. The data are shown in Table 6 with relatively uniform results observed among the different lots. Spike detection (96 to 100 percent) and Change calls (96 to 100 percent), as well as False Change rates (0.16 to 0.26 percent), produced highly consistent results between lots. Lot-to-lot comparisons showed similar performance when analyzed with the same metrics (data not shown).

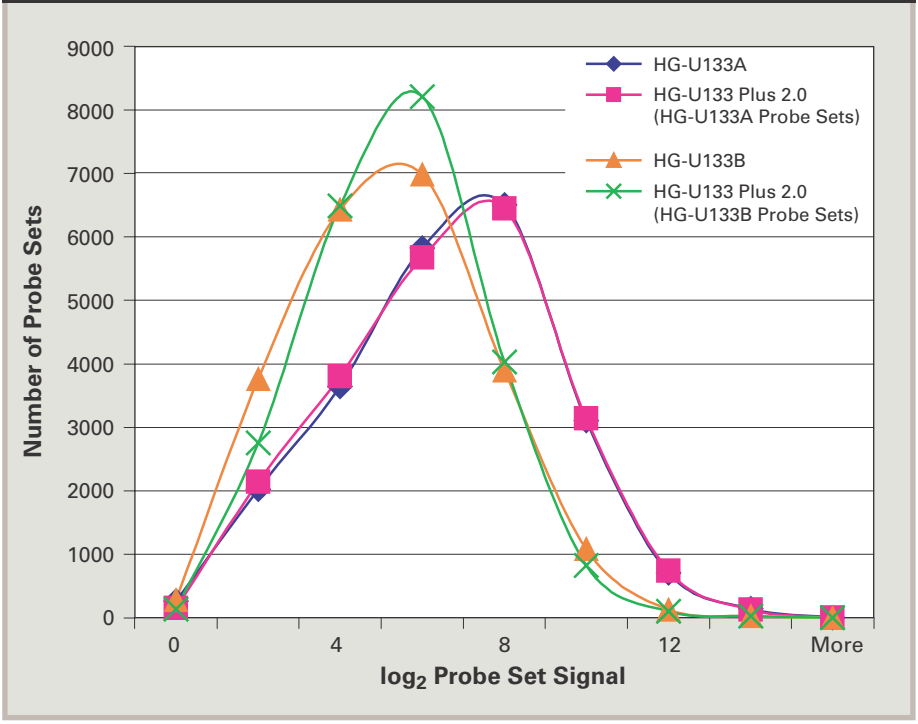
SCALING TO 100 NORMALIZATION CONTROLS

Due to the differences in array content, global scaling strategies were not expected to work well when comparing Signal data from the HG-U133A or HG-U133B Arrays to the new HG-U133 Plus 2.0 Array. (Data demonstrating the drawbacks of a global scaling strategy are presented in the Technical Note, *Performance and Validation of the GeneChip® Human Genome U133 Set*.) An alternative solution to global

Table 6. Array performance from multiple manufacuturing lots. Nine manufacturing lots of HG-U133 Plus 2.0 Arrays were hybridized in triplicate with samples containing 26 human cDNA spikes at varying concentrations as described for the Latin Square experiments. False Change is measured by comparing arrays hybridized with identical sample and is defined as an Increase call with a Signal Log Ratio of ≥ 1 or a Decrease call with a Signal Log Ratio ≤ -1.

Manufacturing Lot	Spike Detection (1.5 pM)	Change Call (3.0 vs. 1.5 pM)	False Change
1	100%	99%	0.21%
2	97%	96%	0.25%
3	96%	96%	0.19%
4	96%	98%	0.26%
5	97%	97%	0.16%
6	97%	96%	0.22%
7	96%	98%	0.19%
8	97%	99%	0.20%
9	96%	100%	0.19%

Figure 6. Probe set Signal distribution after scaling to normalization controls. Sample derived from human prostate was hybridized in triplicate to the GeneChip® HG-U133A, HG-U133B, and HG-U133 Plus 2.0 Arrays. Probe set Signal values were $\log_2$ transformed and then plotted as a frequency histogram, using $\log_2$(Signal) as the bin (x axis), with the number of probe sets in each bin shown on the y axis. Probe sets from the HG-U133 Plus 2.0 Array were plotted separately for those in common with the HG-U133A Array (magenta) and those in common with the HG-U133B Array (green). Distributions for HG-U133A and HG-U133B Arrays are represented by the dark blue and orange curves, respectively.



scaling is to scale the data to the 100 normalization controls common to these three array types. Figure 6 shows data scaled using this strategy.

In this experiment, HG-U133 Plus 2.0 probe sets were first classified as to whether they were derived from the HG-U133A Array or the HG-U133B Array. New probe sets were excluded from this analysis. The data in Figure 6 are presented as a histogram where the y axis represents the number of probe sets per bin and the x axis breaks the data into Signal bins based on the $\log_2$ of the scaled Signal value. As shown in the histogram, the distribution of Signal from the HG-U133A probe sets overlap the Signal produced from the same probe sets on the HG-U133 Plus 2.0 Array. The same can be said for the HG-U133B probe sets, which are plotted in a similar manner. While other strategies may work as well, scaling to the 100 normalization controls creates a similar distribution of Signal values when comparing the same probe sets of different array designs.

Discussion and Summary

The HG-U133 Plus 2.0 and HG-U133A 2.0 Arrays represent the first products in a new generation of expression arrays. The GeneChip Human Genome U133 Plus 2.0 Array provides complete coverage of the Human Genome U133 Set plus 9,921 probe sets representing an additional 6,500 genes for analysis of more than 47,000 transcripts on a single array. The GeneChip Human Genome U133A 2.0 Array represents all of the probe sets present on the HG-U133A Array on a smaller array. The GeneChip Scanner 3000 enabled for high-resolution scanning and GCOS, including the High-Resolution Scanning Patch, are required for scanning and analysis of these new arrays.

Based on the studies shown in this document, the general performance characteristics of the new HG-U133 Plus 2.0 and the HG-U133A 2.0 Arrays are similar to those obtained with the previous HG-U133 Set. However, improved Present call rates were observed with the HG-U133 Plus 2.0 and the HG-U133A 2.0 Arrays, most noticeably among the probe sets in the HG-U133B design. While comparisons between 11-µm and 18-µm array data show a high degree of correlation and concordance, the highest degree of assay reproducibility between samples was observed between arrays with the same feature size.

For the analysis of assay performance, the spike sensitivity experiments were plotted as ROC curves. This type of analysis provides both sensitivity and specificity measurements in a single graph. These graphs show the values produced by the default algorithm parameters as well as the sensitivity-specificity tradeoffs that will be encountered by changing the parameters. As seen in both graphs, high sensitivity values were obtained while maintaining either low false positive (Figure 4) or False Change rates (Figure 5).

Since the completion of the Human Genome, one of the Affymetrix goals was to provide customers a means to interrogate genome-wide transcriptional activity on a single array. The HG-U133 Plus 2.0 Array meets those criteria. In addition to advantages made possible with higher density arrays, both new products – the HG-U133 Plus 2.0 and the HG-U133A 2.0 Arrays – provide customers with the performance and reproducibility that they have come to expect from the Affymetrix Expression Platform.

AFFYMETRIX, INC.

3380 Central Expressway
Santa Clara, CA 95051 USA
Tel: 1-888-DNA-CHIP (1-888-362-2447)
Fax: 1-408-731-5441
sales@affymetrix.com
support@affymetrix.com

AFFYMETRIX UK Ltd

Voyager, Mercury Park,
Wycombe Lane, Wooburn Green,
High Wycombe HP10 0HH
United Kingdom
Tel: +44 (0) 1628 552550
Fax: +44 (0) 1628 552585
saleseurope@affymetrix.com
supporteurope@affymetrix.com

AFFYMETRIX JAPAN K.K.

Mita NN Bldg., 16 F
4-1-23 Shiba, Minato-ku,
Tokyo 108-0014 Japan
Tel: +81-(0)3-5730-8200
Fax: +81-(0)3-5730-8201
salesjapan@affymetrix.com
supportjapan@affymetrix.com

www.affymetrix.com

For research use only.
Not for use in diagnostic procedures.

Part No. 701483 Rev. 2