

Technical Note

Performance Review of the GeneChip® AutoLoader for the Affymetrix GeneChip® Scanner 3000

Designed for use with the GeneChip® Scanner 3000, the GeneChip® AutoLoader provides a temperature-controlled environment for unattended scanning of up to 48 arrays. The AutoLoader includes a 48-array removable carousel, which is maintained at 15°C for all arrays, both pre- and post-scan. For customers who wish to upgrade and include the AutoLoader in their GeneChip applications, trained Affymetrix instrument service engineers can install the AutoLoader and provide introductory training.

When used in conjunction with the GeneChip Scanner 3000, the AutoLoader provides:

- Unattended walk away, automated operation for up to 48 arrays
- Increased ease of use and efficiency
- Sample/process tracking for GLP and future FDA compliance
- Workflow flexibility for batch or continuous processing

This Technical Note describes a set of controlled-comparison experiments which demonstrate the capability of the AutoLoader to provide the benefits of automation without any loss of data quality.

Introduction

The new GeneChip® AutoLoader, used in conjunction with the GeneChip Scanner 3000, provides walk away freedom from scanning, allowing scientists to perform experiments more efficiently and easily. The AutoLoader maintains consistent temperature in a removable 48-array carousel for sample stability and integrity for up to 16 hours at 15-18°C. After arrays are loaded into the carousel, they are each individually scanned by the GeneChip Scanner 3000.

The AutoLoader was designed to provide equivalent data as manual scanning, to support rapid adoption, and to ensure that experimental data can be shared and compared, regardless of the scanning technique used. To this end, data were examined from numerous studies across a set of performance metrics to confirm that the AutoLoader does not significantly affect array results collected on the GeneChip Scanner 3000.

A wide range of selected GeneChip® array designs was tested (Table 1), but the

GeneChip Yeast Genome S98 Array (YG-S98) was used in the majority of the experiments. As the most A/T-rich array in the GeneChip array repertoire, it is also the most sensitive to changes in temperature. Therefore, the YG-S98 Array is a reliable indicator of equal or superior performance across other GeneChip microarrays. Although data are not shown in this paper, the other arrays in Table 1 demonstrated similar performance as the YG-S98 arrays. Arrays with different formats were used to investigate whether the size of the array (larger formats indicate smaller array size) affects the AutoLoader's performance. It was determined that neither array size nor feature size influence AutoLoader performance. Additionally, the data shown in this paper were acquired with an AutoLoader temperature of 15°C. Equivalent data were obtained when the temperature in the AutoLoader carousel was 18°C.

The following pages describe results of controlled-comparison experiments between the AutoLoader and traditional, manual scanning. In all experiments, the GeneChip Scanner 3000 was used. Several experi-

Table 1. Microarrays used in the experimental protocol and their storage temperature in the GeneChip® AutoLoader carousel. GeneChip® brand arrays were used.

Array	Format (arrays/wafer)	Feature Size (µm)	Carousel Templature
Yeast Genome S98	49	24	15°C & 18°C
Human Genome U133A	49	18	15°C & 18°C
Custom Arabidopsis	49	18	15°C & 18°C
Human Genome Focus	100	18	15°C
Rat Genome U34	169	24	15°C
Mapping 10K	49	18	15°C

ments were performed to determine if:

- Array position in the AutoLoader carousel has any effect on array results
- Use of the AutoLoader has any effect on array results compared to manuallyscanned arrays.

The results of these experiments conclusively demonstrate that data generated with the AutoLoader are equivalent to data generated with traditional methods of loading microarrays into the GeneChip Scanner 3000. These findings strongly suggest that the AutoLoader can be rapidly adopted into current laboratory protocols and that data comparisons can be conducted between manually and AutoLoader-scanned arrays.

Experimental Design

The following paragraphs describe the experimental designs used to study the effect of an array's position in the AutoLoader carousel as well as to compare performance between manual scanning and scanning with the AutoLoader.

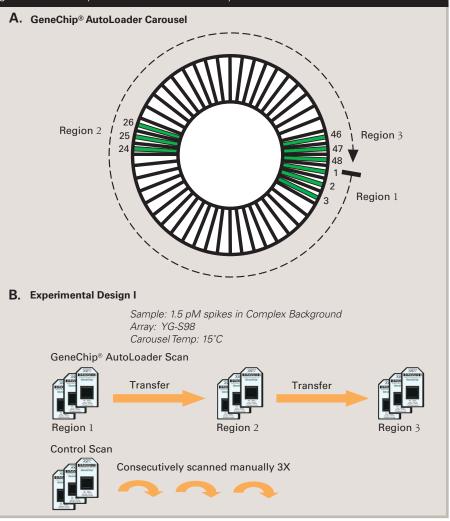
EXPERIMENTAL DESIGN I

In the first experimental protocol, the array's position in the AutoLoader carousel was evaluated to determine if storage at 15°C has any impact on assay performance (Figure 1B). Since it was not practical to compare the performance of each array position to each other, the carousel was divided into three regions as follows (Figure 1A):

- Region 1 (Position 1-3, or P1-P3)
- Region 2 (Position 24-26, or P24-P26)
- Region 3 (Position 46-48, or P46-P48)

Six Yeast Genome S98 Arrays were hybridized with a sample, which consisted of a complex background with eight spike-in transcripts (bioB, bioC, bioD, cre, lys, dap, phe, and thr) at a concentration of 1.5 picomolar (pM). These spikes represent bacterial and phage genes, which were absent in the complex background. The spike-in transcripts were generated by in vitro transcription from vectors contain-

Figure 1. Experimental Protocol and Design. A. Diagram of array position in the GeneChip® AutoLoader carousel. The three-array sets of microarrays are shown in green. The empty slots were filled with previously scanned "dummy" arrays. Numbers represent position (from 1 to 48). **B.** This experiment was used to evaluate the effect of array position in the AutoLoader carousel on results. Three replicate arrays were first loaded into position #1, #2, and #3 (Region 1) of the AutoLoader carousel. After scanning, these three arrays were moved immediately to #24, #25, and #26 (Region 2) followed by #46, #47, and #48 (Region 3) thereafter. A control set of triplicate arrays was scanned manually three consecutive times to assess extent of photo-bleaching of the test arrays. Sample consisted of 1.5 pM spike-in transcripts in a complex cRNA background. GeneChip® Yeast Genome S98 Arrays were used.



ing the cloned DNA sequence. Labeled cRNA transcripts were then quantified, fragmented, and spiked back into labeled, fragmented, complex background cRNA to reconstitute expression of all eight transcripts at 1.5 pM. Large pools of complex sample containing the eight transcripts were created and split among the six arrays.

Three of these arrays were scanned

sequentially, as a set, through each of the three carousel regions (Figure 1B) as follows: First, the three YG-S98 arrays were scanned when they were in position 1-3 (Region 1), again in position 24-26 (Region 2), and finally in position 46-48 (Region 3). Other positions in the carousel were filled with "dummy" arrays to ensure that the time for a test run was equivalent

to the time required to process a carousel with 48 arrays.

The remaining three YG-S98 arrays were used as control arrays and were scanned manually three consecutive times to capture the effects of photo-bleaching*.

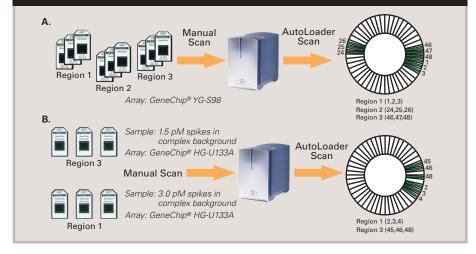
Arrays scanned in Region 1 (P1-P3) of the AutoLoader were compared to the first scans of the control arrays. Arrays scanned in Region 2 (P24-P26) of the AutoLoader were compared to the second scans of the control arrays. Arrays scanned in Region 3 (P46-P48) of the AutoLoader were compared to the third scans of control arrays (Figure 1B). Performance was evaluated with the following metrics:

- Unscaled signal intensity
- Present call percentage
- Detection call accuracy for synthetic spike-in transcripts (bioB, bioC, bioD, cre, lys, dap, phe, and thr) at a concentration of 1.5 picomolar (pM)
- False Change rate
- Detection call concordance

AUTOLOADER VS. MANUAL SCANNING

Two experiments were performed to investigate whether microarrays scanned with the AutoLoader perform equivalently to those scanned manually (Figure 2). The first experiment involved nine GeneChip Yeast Genome S98 arrays in three groups of triplicates (Figure 2A). The sample consisted of a complex yeast background with eight spike-in transcripts (bioB, bioC, bioD, cre, lys, dap, phe, and thr) at a concentration of 1.5 picomolar (pM). These spikes represent bacterial and phage genes, which were absent in the complex background. They were generated by the methods described previously. Each group of triplicates was first scanned manually and then loaded into Region 1 (P1-P3), Region 2 (P24-P26), or Region 3 (P46-P48) to be scanned by the AutoLoader. Data from manually scanned arrays were compared to their AutoLoaderscanned counterparts to assess performance. The carousel was filled with dummy arays to achieve equal run time as a carousel loaded with 48 arrays. Performance was

Figure 2. AutoLoader vs. Manual Scanning. These experiments were used to evaluate whether microarrays scanned with the AutoLoader performed equivalently to those scanned manually. In all experiments, the empty slots were filled with previously scanned "dummy" arrays. A. A total of nine arrays (in three groups of triplicates) were first scanned by conventional manual scanning. The same set of nine arrays was then distributed into positions #1, #2, and #3 (Region 1), positions #24, #25, and #26 (Region 2), and positions #46, #47, and #48 (Region 3), respectively, in the AutoLoader carousel and were scanned. The sample hybridized to the GeneChip® Yeast Genome S98 Arrays consisted of complex cRNA with 1.5 pM spike-in transcripts. The percentage of Present calls between arrays scanned manually and arrays scanned with the AutoLoader were compared. B. Three GeneChip® Human Genome U133A Arrays were hybridized with sample containing 1.5 pM synthetic spike-in transcripts in a complex cRNA background. These arrays were manually scanned and then loaded into positions #2, #3, and #4 (Region 1) in the AutoLoader carousel. Another three HG-U133A arrays hybridized with sample containing 3.0 pM synthetic spike-in transcripts in a complex cRNA background were manually scanned and loaded into positions #45, #46, and #48 (Region 3) in the AutoLoader carousel. Change calls were determined from all nine pair-wise concentration comparisons.



evaluated by comparing the percentage of Present calls between arrays scanned manually and arrays scanned with the AutoLoader.

In the second experiment, GeneChip Human Genome U133A (HG-U133A) arrays were used. Target was prepared from the HeLa cell line. Fifty-three synthetic human spike-in transcripts were added to this complex cRNA at concentrations of either 1.5 or 3.0 pM. (For more information about the sample used, please refer to the Technical Note, The New Affymetrix® GeneChip® Scanner 3000: Seamless Performance Between GeneChip Scanner 3000 and the Affymetrix GeneArray® 2500 Scanner.) Performance was evaluated by Change call accuracy. Three HG-U133A arrays with the spike-in concentration of 1.5 pM were first scanned manually and

then placed in Region 1 of the AutoLoader carousel (P2-P4). Three HG-U133A arrays with the spike-in concentration of 3.0 pM were scanned manually and then placed in Region 3 of the AutoLoader carousel (P45, P46, P48) (Figure 2B). The carousel was filled with dummy arays to achieve equal run time as a carousel loaded with 48 arrays. Average Change call accuracy was calculated for all possible pair-wise comparisons of the three replicates at two concentrations, resulting in nine comparison analyses (Table 2).

*These photo-bleaching control arrays were not meant to serve as controls for the time delay that occurs as a function of an array's position in the AutoLoader. Instead, the photo-bleaching control arrays were used as a control only for the photo-bleaching that may be a result of scanning each array three consecutive times, which is not typically included in the standard GeneChip scanning protocol.

Performance Data & Analysis

In all experiments, the GeneChip Scanner 3000 was used. Performance was assessed on the basis of both single array and comparison analyses generated with Affymetrix® Microarray Suite (MAS) 5.0 software, which contains the same algorithms as the newly released GeneChip® Operating Software (GCOS). Analysis output files (.CHP files) were created at Affymetrix by globally scaling all probe sets on image data files (.DAT files) to a target intensity of 250, except when unscaled signal intensity was evaluated.

Data from each method of scanning (manual vs AutoLoader) were tested for statistical significance by applying paired Student's t-tests. The null hypothesis was defined as follows: manual scans significantly exceed the AutoLoader performance at a p>0.95 confidence level. Therefore, p-values greater than or equal to 0.95 were interpreted as supporting the null hypothesis (indicating significantly better performance for the manual scanning method); p-values approaching 0.5 were interpreted as equivalent performance; and p-values less than or equal to 0.05 were interpreted as contradicting the null hypothesis (indicating significantly better performance for AutoLoader scans).

Results & Discussion

GeneChip expression array data comparing the performance of the AutoLoader to manual scanning methods indicated that there is no significant performance difference when the AutoLoader is used. This conclusion was reached by examining data from numerous experiments and evaluating performance metrics described below. Both methods of scanning produce high quality data for gene expression analysis with GeneChip microarrays. However, as the AutoLoader enables walk away freedom, it offers high-quality data, and is particularly amenable to high-throughput environments.

RESULTS SUMMARY:

Evaluation of Position in Carousel on Performance

In order to assess if performance changes as a function of location in the AutoLoader carousel, the experiment outlined in Figure 1 was performed and the following metrics were evaluated:

- Unscaled signal intensity
- Present call percentage
- Detection call accuracy for synthetic spike-in transcripts (*bioB*, *bioC*, *bioD*, cre, *lys*, *dap*, *phe*, and *thr*) at a concentration of 1.5 picomolar (pM)
- False Change rate
- Detection call concordance

UNSCALED SIGNAL INTENSITY

Sequential scanning of arrays in each of the three regions of the AutoLoader results in different time exposures to 15°C storage temperature because the final arrays to be scanned are left in the carousel for the greatest length of time. To assess whether there is any appreciable degradation in array performance due to storage at 15°C, relative signal values were compared as a function of microarray position in the carousel. Arrays in Region 1 were scanned first (and had the least exposure to the 15°C

storage conditions), while arrays in Region 3 were scanned last (and had greatest length of exposure to 15°C storage conditions). In order to evaluate the effect of carousel position on signal, unscaled signal values were used.

All signal values generated from arrays in each region were averaged and compared as described in the following sentences. First, the signals for the set of three arrays from Region 1 were averaged and considered to be 100 percent. The signal for the three-array sets from Regions 2 and 3 were then normalized relative to this 100 percent value. The same data processes were applied to the photo-bleaching controls.

After the normalized relative signal values were obtained, a ratio of relative signal value (of the AutoLoader-scanned arrays compared to photo-bleaching controls) was calculated for Region 2 and Region 3. These two values were 91 percent and 88 percent of the photo-bleaching controls respectively. This can be seen in Figure 3 as a slight, yet noticeable, signal decrease in Regions 2 and 3 compared to Region 1. However, when other parameters were evaluated, this minor decrease in signal had

Figure 3. Relative signal values corresponding to the positions of arrays in the GeneChip® AutoLoader carousel. Data points were averaged from the three replicate arrays in each carousel region and plotted with their corresponding data from the control scans. The orange line indicates photo-bleaching controls, which were manually scanned. The dark blue line indicates the AutoLoader-scanned arrays. The y axis indicates signal values relative to the signal value of Region 1.

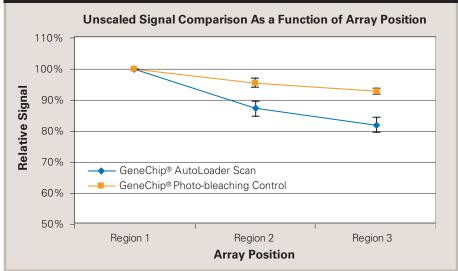
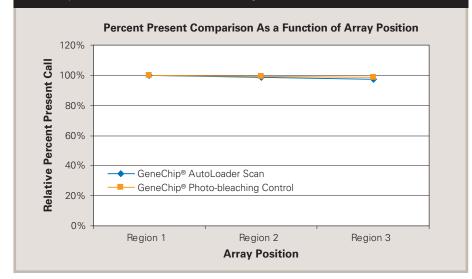


Figure 4. Relative Present call percentage corresponding to array position relative to Region 1 in the GeneChip® AutoLoader carousel. Relative Present call percentage for each array was calculated with Affymetrix® Microarray Suite 5.0 analysis software. Values for the three replicates were averaged to obtain the region-specific percent Present call. Percent Present calls were generated for Regions 2 and 3 relative to Region 1. The orange line indicates photo-bleaching controls, which were manually scanned. The dark blue line indicates the AutoLoader-scanned test arrays. The y axis indicates percent Present calls relative to Region 1.



minimal impact on overall array performance, as described later in this paper.

PRESENT CALL PERCENTAGE

In general, the number of Present calls in replicate samples within an experiment should be similar to each other as each replicate represents identical target sample. To calculate the consistency of the GeneChip Scanner 3000 when used with the AutoLoader, the percent Present call was averaged for each three-array set and compared. As demonstrated in Figure 4, there was a barely discernable decrease in percent Present calls for arrays placed further down the scanning queue (i.e., Region 1 vs. Region 2 vs. Region 3). At most, there appeared to be an approximately two percent decrease in terms of Present calls between arrays in Region 3 and Region 1 (after subtracting the effect of photobleaching). This rather minute difference was concluded to be insubstantial and was not corroborated by other metrics such as False Change rate, Detection call accuracy for spikes, or Detection call concordance.

Furthermore, the differences observed between regions are comparable to typical variation observed between replicate arrays.

1.5 PM SPIKE DETECTION CALL ACCURACY

The use of target spikes at low concentrations (1.5 pM) in a complex biological background provides a calibrated measure of instrument performance at the limits of detection of the scanning system. For example, transcripts at 1.5 pM correlate to a frequency of one transcript in 100,000 and are considered a reliable measurement of GeneChip instrument system sensitivity.

Detection call accuracy can be ascertained by hybridizing microarrays to a complex cRNA background with a known concentration of spike-in transcripts and determining the percentage of accurate Present calls. For each of the three-array sets in the GeneChip AutoLoader, as well as the control arrays, eight cRNA spikes were added to the complex cRNA target at a 1.5 pM concentration.

Percent Present calls for the spiked transcripts were averaged for each three-array set and compared to determine the effect position in the carousel had on Detection call accuracy of arrays scanned on the GeneChip 3000 Scanner and stored in the AutoLoader. Figure 5 shows that spike

Figure 5. Effect of array position in the GeneChip® AutoLoader carousel on Detection call accuracy for spikes at a 1.5 pM concentration. Each data point is a result of averaged spike Present call accuracies for three replicate arrays. The orange line indicates photo-bleaching controls, which were manually scanned. Photo-bleaching controls were plotted with their corresponding data from the AutoLoader. The dark blue line indicates the AutoLoader-scanned arrays. The y axis indicates the relative accuracy of spike percent Present call accuracy relative to Region 1.

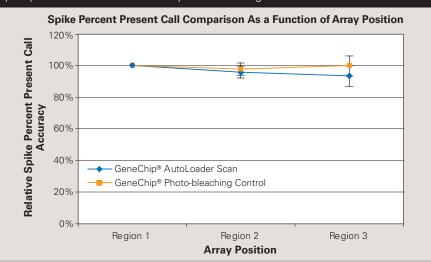
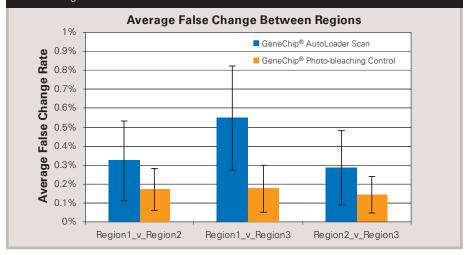


Figure 6. False Change rate corresponding to array position in the GeneChip® AutoLoader carousel. Each vertical bar represents the average of nine false change values derived from three-by-three array comparisons between two carousel regions. Error bars represent the standard deviation. The x axis represents comparisons between three distinct regions of the carousel, namely, Region 1 vs. Region 2, Region 1 vs. Region 3, and Region 2 vs. Region 3. The y axis corresponds to average False Change rate.



Detection call accuracy was consistent, regardless of an array's position in the AutoLoader carousel. Student's *t*-tests confirmed this conclusion (data not shown).

FALSE CHANGE RATE

False Change measures the reproducibility of replicate arrays within an experiment. It is defined as the frequency of the sum of Increase calls greater than or equal to the Signal Log Ratio of 1 and Decrease calls less than or equal to a Signal Log Ratio of -1 (as determined by Affymetrix Microarray Suite analysis software). Therefore, to determine how the AutoLoader affects the reproducibility of data collected on the GeneChip Scanner 3000, False Change was analyzed for each three-array set in each region of the AutoLoader and compared. Affymetrix specifications dictate that the False Change rate between replicate arrays must not be greater than one percent and should be comparable to the value derived from traditional manual scans (photobleaching control).

Figure 6 shows the three possible region-to-region, pair-wise comparisons for both AutoLoader and control methods. For a given region-to-region comparison

there were nine possible pair-wise comparisons involving three replicate arrays from one region and three from the other, giving rise to nine False Change values for each particular pair. Vertical bars represent the average False Change values for each

pair-wise, region-to-region comparison. Error bars represent one standard deviation. The data clearly indicate that the False Change rates of AutoLoader scans are consistent across the three carousel regions and well below the one percent False Change specification. The same experiments were repeated with 18°C carousel storage temperatures and similar results were obtained (data not shown).

DETECTION CALL CONCORDANCE

Detection call concordance between replicate arrays is often used with False Change rates to examine consistency. In this study, analysis involved three possible region-to-region, pair-wise comparisons (Region 1 vs. Region 2, Region 1 vs. Region 3, and Region 2 vs. Region 3) for a total of nine concordance values for a given comparison. Figure 7 summarizes the average Detection call concordance for each respective region-to-region comparison, with the AutoLoader scans shown in blue and manual control scans in orange.

The data clearly indicate the fact that arrays scanned via the AutoLoader maintained high call concordance similar to manually scanned arrays (~96 percent in

Figure 7. Detection call concordance corresponding to array position in the GeneChip® AutoLoader carousel. Each data point was averaged from nine Detection call concordance values derived from three-by-three microarray comparisons between two carousel regions. The x axis represents comparisons between three distinct regions of the carousel, namely, Region 1 vs. Region 2, Region 1 vs. Region 3, and Region 2 vs. Region 3. The y axis corresponds to Detection call concordance between regions.

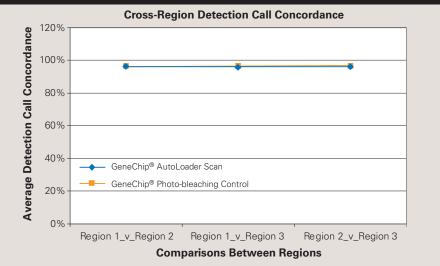
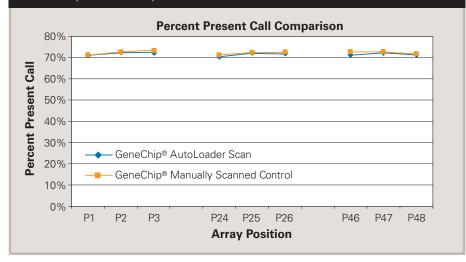


Figure 8. Direct Present call comparison between manual control scans and GeneChip® AutoLoader scans. The x axis represents the carousel position for the nine microarrays scanned. The y axis represents the percent Present call for each array. The blue line represents the GeneChip AutoLoader scan while the orange line represents the manually scanned arrays.



both cases). There was virtually no difference in call concordance values - or consistency – between arrays loaded in Region 1 vs. Region 2, Region 1 vs. Region 3, or Region 2 vs. Region 3 as noted by the flatness of the curve in Figure 7.

RESULTS SUMMARY:

AutoLoader vs. Manual Scanning

The following experiments were performed to determine if arrays scanned with the AutoLoader performed equivalently to arrays scanned manually.

PRESENT CALL COMPARISON

This experiment evaluated the Auto-Loader's effect on data accuracy by comparing gene expression results for YG-S98 arrays scanned with the AutoLoader to the gene expression results for the same arrays generated from manually scanned arrays (Figure 2A). Figure 8 shows the Present call comparisons between the manual scans of three arrays in triplicate and AutoLoader scans of the same arrays, which were inserted in Regions 1, 2, and 3 of the carousel. Present calls displayed high concordance and were not affected by scanning method.

CHANGE CALL ACCURACY FOR SPIKED TRANSCRIPTS

Change call accuracy is a qualitative measurement used to indicate the difference in

the transcript concentrations between a baseline array and an experiment array, and is, therefore, a measure of array sensitivity. The Change call accuracy can be determined by incorporating various, known concentrations of spike-in transcripts known to be absent in the complex cRNA background sample and comparing the resulting probe-level analyses. To determine if the AutoLoader impacts array sensitivity, comparison experiments were conducted with two different concentrations of spike-in transcript against a complex cRNA background (Figure 2B).

Table 2 summarizes the results of this experiment which includes nine Change calls obtained by comparing HG-U133A arrays with two spike-in concentrations (1.5 pM vs. 3.0 pM) of 53 human transcripts. The Change call accuracy was calculated for manually scanned as well as for AutoLoader-scanned arrays. A paired Student's *t*-test indicated that there was no statistical significance between the two methods of scanning.

Table 2. Equivalent Change call accuracy of GeneChip® AutoLoader-scanned arrays vs. manually scanned arrays. Nine pair-wise comparisons were conducted between the three arrays with spike-in transcripts at a concentration of 1.5 pM and three arrays with spike-in transcripts at a concentration of 3.0 pM. Change call accuracies were calculated for manually and AutoLoader-scanned arrays.

Array Comparisons (1.5 pM vs. 3.0 pM spikes)	GeneChip® AutoLoader Scan	Manual Scan
P2_v_P45	98.1%	96.2%
P2_v_P46	100.0%	98.1%
P2_v_P48	98.1%	98.1%
P3_v_P45	98.1%	100.0%
P3_v_P46	96.2%	98.1%
P3_v_P48	96.2%	98.1%
P4_v_P45	98.1%	96.2%
P4_v_P46	96.2%	98.1%
P4_v_P48	98.1%	98.1%
Standard Deviation	1.3%	1.1%
Average	97.7%	97.9%

Summary

The results presented all support the hypothesis that there is no significant degradation in the quality of array data as a result of the microarray's position in the GeneChip® AutoLoader. To evaluate the effect of the AutoLoader on data accuracy, gene expression results from GeneChip arrays scanned with the AutoLoader were compared to the gene expression results generated from manually scanned arrays.

There was no statistically significant difference between manually scanned and AutoLoader-scanned arrays, which indicates that using the AutoLoader has no effect on the accuracy of gene expression data. Additionally, this finding suggests that data generated from manual scans can be compared to arrays scanned with the AutoLoader.

Both methods produce high-quality data for gene expression analysis with GeneChip microarrays. However, as the AutoLoader enables walk away freedom, the GeneChip AutoLoader provides a pivotal link between high-quality array data and high-throughput automated array analysis.

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

www.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

For research use only.

Not for use in diagnostic procedures.

Part No. 701515 Rev. 1

©2004 Affymetrix, Inc. All rights reserved. Affymetrix®, GeneChip®, ✓✓ * * * * , ✓ * , ► ✓ *