AFFYMETRIX®

Technical Note

Manufacturing Quality Control and Validation Studies of GeneChip® Arrays

Quality control and validation are integral parts in the Affymetrix GeneChip® array production process. Numerous manufacturing and validation tests are performed to ensure these processes meet quality control and validation standards. Quality control in manufacturing monitors three areas—design, synthesis and signal intensity—while validation assesses functional performance, including sensitivity and reproducibility. This document summarizes the extensive tests that are performed to ensure a high level of quality in GeneChip arrays.

Manufacturing Quality Control

Our general manufacturing quality control strategy verifies three aspects:

- Design
- Synthesis
- Signal Intensity

To support this strategy, we adapted knowledge gained from two mature industries—the oligonucleotide synthesis industry and the semiconductor industry—to influence the standardization of our verification process.

Specifically, we use many reagents within our chemistry process that are also used in standard oligonucleotide synthesis. The main difference in these two processes is the use of light versus acid to deprotect and build probes. In regard to the semiconductor industry, we draw from their existing processes and equipment for substrate preparation, photolithography, and packaging applications, and adapt them for our own uses.

Design Verification

After an extensive sequence selection process, a set of photolithographic masks are designed to represent the information content. Within this mask design process, automated software tests verify the design by building the array *in silico*. Through this virtual synthesis, these software tools verify, and thus ensure, the correct probe sequences will be synthesized in the correct (x, y) location within the array.

Please refer to the *Array Design for the GeneChip Human Genome U133 Set* Technical Note for additional information regarding the sequence selection and probe selection steps of the design process.

Synthesis Verification

Another step used to ensure quality within our manufacturing process is the use of a manufacturing execution software system adapted from the semiconductor industry. One aspect of the system verifies that the correct reagent delivery occurs at the appropriate time, with each step of the synthesis process interactively tracked.

To enable this tracking process, each mask in a design set has a unique identifier. Optical character recognition (OCR) is used to verify that the correct mask and wafer are used at the appropriate time. The software also tracks and validates each mask identifier so the masks cannot be used out of sequence. This system ensures excellent control of our process and that the correct synthesis of the product occurs.

Combinatorial Chemistry Strategy

Our challenge in the quality testing of array synthesis is ensuring the oligonucleotide probes on the array are synthesized correctly.

To address this challenge, we have developed strategies that involve the manufacturing synthesis process in combination with specially designed control probes. These probes are synthesized on the array in conjunction with the product probes, ensuring that each step in the manufacturing process is correct.

Figure 1. Synthesis Confirmation.

Mask Cycles	1	2	3	4	5	6	7	8	Probes Synthesized
	A	C	G	T	A	C	G	T	
Probe 1	A	C	-	T	-	C			A C T C
Probe 2			G	-	A	-	G	T	G A G T
Probe 3	A	-	G	-	A	-	G		A G A G
Probe 4		C	-	T	-	C	-	T	C T C T

Because it is not practical to test each oligonucleotide individually, we use these control probes in combination with the power of combinatorial chemistry to our advantage. Figure 1 illustrates an example of an array that is composed of only four probes with the sequences listed. We build these four sequences using eight separate chemistry cycles and eight masks. Probe 1 is built in cycles 1, 2, 4, and 6, while probe 2 is built in cycles 3, 5, 7, and 8.

To confirm the four probes are synthesized accurately, we can analyze probes 1 and 2, which, together, cover all eight cycles. (Conversely, we could analyze probes 3 and 4 to cover the eight cycles.) Analysis of either set of probes represents all the cycles used to make the array; therefore, we do not need to analyze all the probes on the array to verify their correct synthesis. Using this strategy, we validate each cycle in the synthesis of our commercially available arrays.

In practice, synthesis is verified by using a more complex version of this simplified example. Using a panel of control probes that represent each chemistry cycle, we can check for chemistry and photolithography errors, as demonstrated in Figure 2. Here, each column represents a chemistry step. Wafer A represents an acceptable synthesis in which there are no missing probe cells, and the probe cell pattern is aligned correctly. Wafer B illustrates an unacceptable synthesis in which probe cells are missing, as indicated. The probe cell location identifies the errant photolithography and chemistry cycles.

Signal Verification: Using Hybridization Controls

A second panel of control probes is built within each array design as an additional quality test. This test verifies that the array yields sufficient signal levels, as set by the product specification.

An example of this strategy is demonstrated in the GeneChip Human Genome U133 (HG-U133) Set. The HG-U133 Set contains a panel of probe sets that represent hybridization control genes consisting of *bioB*, *bioC*, *bioD*, and *cre*. *BioB*, *bioC*, and *bioD* are genes in the biotin synthesis pathway of *E. coli*, and *cre* is the recombinase gene from P1 bacteriophage. These control probe sets are used to verify that the array yields a sufficient signal to meet the minimal threshold for product requirement specifications.

SECTION SUMMARY

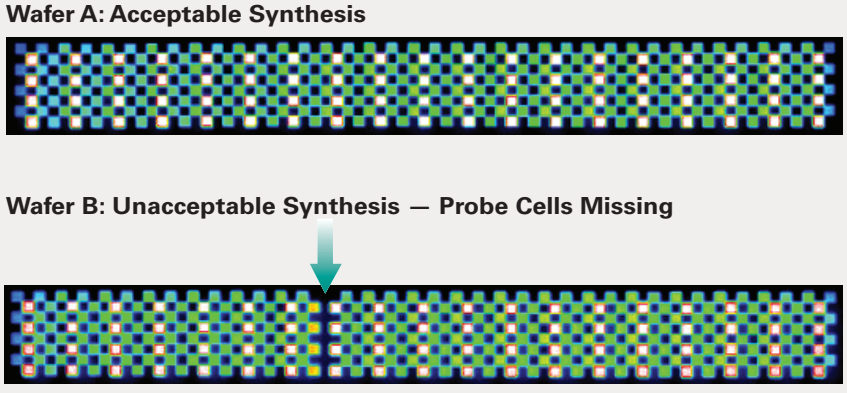
As described previously, quality is ensured throughout the entire manufacturing process by verifying design, synthesis, and the signal intensity. The array mask set designs are verified by an automated design check process. In-process software controls and post-process array testing verify synthesis. The signal intensity is verified by assay hybridization controls.

Accuracy of the sequence and efficiency of the nucleotide synthesis reaction are, thus, ensured as we test for both of these items using redundant testing strategies as outlined. Failure of either the sequence accuracy or the synthesis reaction will be apparent in any one of these quality assays, and the point of failure can be determined by analysis of the hybridization patterns in these tests.

Validation Studies

In addition to making the array properly, it is also necessary to test the array for functional performance. We do this by specifically testing the sensitivity and reproducibility of each design. As an example, we analyzed the HG-U133 Set by using the Affymetrix® Microarray Suite 5.0 (MAS 5.0) software. The analysis yielded a detection sensitivity level of one transcript in 100,000, with a false change rate of 1% or less. (Other arrays may have slightly different false change criteria, which can be found in the individual product specifications.)

Figure 2. Synthesis Verification Results.



Sensitivity

Pre-labeled transcripts, spiked into a complex background at varying concentrations, are typically used to demonstrate array sensitivity. In these experiments, sensitivity was measured by quantitative metrics from both single array and comparison analyses.

Again using the HG-U133 Set as an example, sensitivity was assessed by measuring the Detection calls of pre-labeled human transcripts spiked at 1.5 pM in 50 µg/mL of complex target. This concentration corresponds to approximately one transcript in 100,000, or 3.5 copies per cell.

Sensitivity in a comparison analysis was measured by Change calls of spikes at 3.0 pM in a complex target, when compared to a sample with spikes at 1.5 pM. This difference measures a two-fold change (Signal Log Ratio of 1) for transcripts present at low abundance.

Using the MAS 5.0 software, the detection criterion for spike probe sets is at least 70% or greater Present calls for 1.5 pM spikes, which correlates to the sensitivity specification of one transcript in 100,000.

In a comparison analysis, the criterion is at least 80% or greater Increase calls for a sample with 3.0 pM spikes as the experimental sample, compared to a baseline sample with 1.5 pM spikes. As both of these tests are performed at low concentrations in a complex background of labeled cRNA, we are testing the ability to discriminate true signal from background system noise.

In the following validation experiment for the HG-U133 Set, the sensitivity detection analysis used a sample with 54 HG-U133A or 48 HG-U133B human transcripts, and four bacterial transcripts spiked at 1.5 pM. The nine replicates (triplicates on three lots each of HG-U133A and HG-U133B) with 1.5 pM spikes were evaluated for Detection calls.

Using the 11 probe pair analysis default settings in the MAS 5.0 software, an RNA spike probe set is defined as detected if it receives a Detection call of Present. The number of detected human spike probe sets was averaged over the three replicates on each lot. The percentage of detected human spike probe sets was then calculated for each array type and lot.

Table 1 summarizes these results. For Present calls, the mean detection is based on a total of nine replicates (three replicates each on three lots of arrays). For Increase calls, the mean is based on 27 comparison analyses (nine comparisons of three 3.0 pM replicates versus three 1.5 pM replicates per lot).

Table 1. Sensitivity.

	HG-U133A	HG-U133B
Mean % Present Calls	79%	77%
Standard Deviation Present Calls	8%	3%
Mean % Increase Calls	94%	87%
Standard Deviation Increase Calls	3%	7%

Table 2. False Change.

	Mean % False Change	Standard Deviation
Adult Brain	0.18%	0.05%
Adrenal	0.14%	0.02%
Kidney	0.14%	0.003%

Reproducibility: False Change and Concordance

FALSE CHANGE

Measuring the false change rate and the concordance of Detection calls tests the reproducibility of the arrays. A false change is defined in a comparison analysis between replicate samples as Increases with a Signal Log Ratio value greater than or equal to (≥)1, or Decreases with a Signal Log Ratio value less than or equal to (£) -1 (a two-fold or greater change).

For the HG-U133 Set, the acceptance criterion for false change was 1% or less for comparison analyses of technical replicates. The percent false change is calculated as follows: (number of false changes/total number of non-control probe sets) x 100%. The percent false change was generated from experiments using labeled human cRNA carried out on HG-U133A and HG-U133B arrays as follows. For a given set of three arrays hybridized to an identical target, all three pair-wise Change call analyses were generated: A versus B; A versus C; and B versus C. The percent false change was generated for each analysis. The average of three percent false change values was used to determine if the false change acceptance criteria were met for a given sample and array lot.

Representative data from human tissue samples are shown in Table 2. False change percentages are based on triplicate comparison analyses.

CONCORDANCE

The false change rate is a measure of reproducibility in comparison analyses. Detection call concordance is another way to evaluate the array reproducibility using single chip analyses. The percentage of concordant Detection calls for two arrays is the sum of the number of concordant Present (P) and concordant Absent (A) or Marginal (M) calls.

Table 3 represents the average detection call (Present and Absent) results from triplicate hybridizations of samples derived from various starting material. Three identical sample cocktails (represented as 1, 2, and 3) are hybridized to arrays of the same type. The Detection calls are compared in a pair-wise fashion as follows: 1 versus 2; 1 versus 3; and 2 versus 3.

Table 1. Sensitivity.

		HG-U133A		HG-U133B	
Tissue		**Number**	**Percent**	**Number**	**Percent**
Adult Brain	Concordant P	8208	36%	4858	21%
	Concordant A/M	12,775	57%	15,976	70%
	Total Concordant Calls	**20,983**	**93%**	**20,834**	**91%**
Adult Heart	Concordant P	3512	16%	1496	7%
	Concordant A/M	17,476	77%	19,638	86%
	Total Concordant Calls	**20,988**	**93%**	**21,134**	**93%**
Pancreas	Concordant P	6556	29%	3266	14%
	Concordant A/M	14,473	64%	17,642	77%
	Total Concordant Calls	**21,029**	**93%**	**20,908**	**91%**
Placenta	Concordant P	8432	37%	4283	19%
	Concordant A/M	12,576	56%	16,671	73%
	Total Concordant Calls	**21,008**	**93%**	**20,954**	**92%**

The results for each tissue represent the average of triplicate hybridizations. The percentage of concordant Detection calls (Table 3) ranges from 91% to 93% for the experiments with HG-U133A and HG-U133B arrays.

For example, the average concordance for an adult brain sample hybridized to HG-U133A and HG-U133B arrays was 93% and 91%, respectively. Overall, while the number of concordant Present calls varies by tissue and array type, the overall percentage of concordant calls is consistent.

Additional information regarding validation of the HG-U133 Set is available in Technical Note 2: *Performance and Validation of the GeneChip Human Genome U133 Set*.

Summary

A comprehensive set of manufacturing and validation tests are performed throughout these processes.

In manufacturing, quality control is monitored in three areas:

- ***Design*** is monitored through automated software tests.
- ***Synthesis*** uses a manufacturing execution software system for tracking purposes and specific probe synthesis sequences.
- ***Signal intensity*** is tested with hybridization controls.

In validation, functional performance is assessed by testing sensitivity and reproducibility as demonstrated with the HG-U133 Set.

- ***Sensitivity*** was measured both by detection of 1.5 pM human clone spikes and as a two-fold difference in Change calls between 1.5 pM and 3 pM human clone spikes.
- False change and concordance measure ***reproducibility.*** An acceptable criterion of a false change rate of 1% or less for comparison analyses of replicates was produced.

In conclusion, our manufacturing and validation testing ensure a high level of quality in Affymetrix' GeneChip arrays.

AFFYMETRIX, INC.

3420 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
UK and Others Tel: +44 (0) 1628 552550
France Tel: 0800919505
Germany Tel: 01803001334
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 **Please visit our web site for international distributor contact information.**

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

Part Number 701309 Rev. 3