2

Expression Profiling Using Affymetrix GeneChip® Probe Arrays

Martina Schinke-Braun and Jennifer A. Couget

Summary

Large-scale microarray expression profiling studies have helped us to understand basic biological processes and to classify and predict the prognosis of cancers; they have also accelerated the identification of new drug targets. Affymetrix GeneChip® probe arrays are high-density oligonucleotide microarrays that are available for many prokaryotic and eukaryotic species. Affymetrix human and mouse whole-genome microarrays analyze the expression level of up to 47,000 transcripts and variants. Each transcript is measured by 11 probe pairs, which consist of a perfect match 25mer oligonucleotide (PM) and a 25mer mismatch oligonucleotide (MM) that contains a single base pair mismatch in the central position. The PM/MM design is used for identification and subtraction of nonspecific hybridization and background signals.

Advantages of Affymetrix GeneChip arrays include highly standardized array fabrication, as well as standardized target preparation, hybridization, and processing protocols, resulting in low technical variability and good reproducibility between experiments. This chapter describes the standard assay procedures for isolating total RNA from heart tissue, generating a biotin-labeled target for expression analysis, processing of Affymetrix GeneChip probe arrays using the Affymetrix instrument system, and quality control measures.

Key Words: Affymetrix; GeneChip; expression profiling; RNA amplification; RNA quality; Agilent Bioanalyzer; in vitro transcription; heart.

1. Introduction

Although many microarray technology platforms are available, Affymetrix GeneChip microarrays are still the most commonly used; Affymetrix has been a pioneer in the microarray field, and its technological strength has been consistently ranked among the top three in the pharmaceutical and biotech sector by MIT's *Technology Review*. Affymetrix GeneChip microarrays are also

From: *Methods in Molecular Biology, vol. 366: Cardiac Gene Expression: Methods and Protocols*
Edited by: J. Zhang and G. Rokosh © Humana Press Inc., Totowa, NJ

superior to other technology platforms because their technology is well established and standardized protocols and an automated instrument setup for microarray processing are used.

The generation of gene expression microarray data using Affymetrix GeneChip probe arrays follows a simple procedure consisting of six major steps ***(1)***. The single most important step in ensuring a successful GeneChip experiment is the preparation of a clean, intact RNA sample. After the RNA sample passes a quality assessment, the labeling process is initiated ***(2)***. Briefly, the first step of the labeling procedure is the synthesis of double-stranded cDNA from the RNA sample using reverse transcriptase and an oligo-dT primer ***(3)***. Next, the cDNA serves as a template in an in vitro transcription (IVT) reaction that produces amplified amounts of biotin-labeled antisense mRNA, which is also referred to as labeled cRNA or the microarray target. Prior to hybridization, the cRNA is fragmented to 50- to 200-base fragments using heat and Mg^{2+}, which facilitates efficient hybridization ***(4)***. The cRNA is added to the hybridization cocktail, which contains salts, blocking agents, and bacterial RNA spike-in controls. This cocktail is injected into the GeneChip array hybridization chamber and hybridized at 45°C for 16 h ***(5)***. After hybridization, the GeneChip array is subjected to a series of washing and staining steps involving a fluorescent molecule that binds to biotin and a signal amplification step ***(6)***. The GeneChip array is then scanned with a confocal laser scanner, and the image of the distribution pattern of fluorescent signals on the microarray is recorded. Control measures are implemented at each step to ensure good-quality data.

Although processing of Affymetrix GeneChip probe arrays is highly standardized and automated, substantial variation has been observed owing to factors such as RNA isolation batches, hybridization day, wash and reagent batches, operator, and lot-to-lot array variation ***(1)***. The effects of the processing may be even larger than the biological effects being studied. To keep variability to a minimum, it is advantageous to label and process the samples for a given experiment at the same time. If the number of samples in an experiment is too large, care must be taken not to confound the treatments with the processing. However, if sample preparation and processing batches are being tracked and incorporated in the design and analysis of microarray experiments, their impact on the data can be removed using statistical techniques such as analysis of variance.

The protocols in this chapter that pertain to sample labeling and Affymetrix GeneChip microarray processing are condensed versions excerpted from the GeneChip Expression Analysis Technical Manual (with permission from Affymetrix). The purpose is to give an overview of the process and point out some of the important tricks and details. For the detailed protocols, it is strongly recommended to refer to the appropriate Affymetrix manuals.

2. Materials

2.1. Heart Tissue Isolation

1. Phosphate-buffered saline (PBS), pH 7.4, at 4°C.
2. 70% Ethanol.
3. Sterile gauze.
4. Liquid nitrogen.

2.2. RNA Extraction from Heart Tissue

1. Power homogenizer (Omni PCR Tissue Homogenizing Kits, Omni International, or similar) with 7-mm saw-toothed generator probe.
2. Small stainless-steel spatula.
3. Filter plugged, sterile, RNase-free pipet tips of various volumes.
4. Disposable polystyrene sterile 75-mm transparent loose-cap tubes (Fisher Scientific, cat. no. 4-956-30).
5. Sterile, RNase-free Eppendorf tubes.
6. TRIZOL reagent (Invitrogen, cat. no. 15596-026).
7. Chloroform.
8. Isopropanol.
9. Ethanol (100%).
10. Diethyl pyrocarbonate (DEPC)-treated, autoclaved water (Ambion, cat. no. 9920).
11. Phenol/chloroform/isoamyl alcohol (25:24:1), pH 7.9 (Ambion, cat. no. 9730).
12. Isoamyl alcohol/chloroform (1:24).
13. PBS, pH 7.4, at 4°C.
14. RNeasy MiniKit (Qiagen, cat. no. 74104); contains 50 RNeasy mini spin columns, collection tubes, and RNase-free reagents and buffers.

2.3. RNA Quality Control and Quantitation Using the Agilent Bioanalyzer

1. Bioanalyzer RNA Nano Kit (Agilent, cat. no. 5064-8229). Includes 25 chips and reagents (gel, dye, filter columns, RNA Nano marker).
2. RNase Zap (Ambion).
3. DEPC water (Ambion).
4. Agilent 2100 Bioanalyzer.
5. Chip priming station (Agilent, cat. no. 5065-4401).
6. Vortexer (e.g., IKA model MS1) with vortex mixer adapter (Agilent, cat. no. 5022-2190).

2.4. RNA Amplification and Labeling Using Affymetrix One-Cycle Labeling Reagents

For steps 1–5, of **Subsection 3.5.**

1. One-Cycle Target Labeling and Control Reagents (Affymetrix, cat. no. 900493). This kit contains all required labeling and control reagents to perform 30 one-cycle labeling reactions. All components (IVT Labeling Kit, One-Cycle cDNA Synthesis Kit, GeneChip Sample Cleanup Module, Poly-A RNA Control Kit,

and Hybridization Control Kit) can also be purchased separately. The Oligo-dT primer (50 μ*M*, high-performance liquid chromatography [HPLC] purified, Affymetrix, cat. no. 900375) is also included and has the sequence:

5' GGCCAGTGAATTGTAATACGACTCACTATAGGGAGGCGG-$(dT)_{24}$ 3'.

2. Absolute ethanol (stored at –20°C for RNA precipitation; stored at room temperature for use with the GeneChip Sample Cleanup Module).
3. 80% ethanol (stored at –20°C for RNA precipitation; store ethanol at room temperature for use with the GeneChip Sample Cleanup Module).
4. 3 *M* Sodium acetate (NaOAc; Sigma-Aldrich, cat. no. S7899 or similar).
5. 1 *N* NaOH.
6. 1 *N* HCl.

2.4.1. Fragmentation of Labeled cRNA

1. 5X Fragmentation Buffer: 200 m*M* Tris-acetate, pH 8.1, 500 m*M* KOAc, 150 m*M* MgOAc in DEPC-treated water; filter through a 0.2-μm filter, aliquot, and store at room temperature. This buffer is included in the Affymetrix GeneChip Clean-up Module (Affymetrix, cat. no. 900371).
2. DEPC-water (Ambion, cat. no. 9920 or similar).
3. Labeled cRNA from **Subheading 2.3.**, **item 5**.
4. 0.5 mL RNase-free tubes (Ambion, cat. no. 12250 or similar).

2.5. Controls

2.5.1. Poly-A RNA Spike-In Controls

1. 1 Poly-A RNA Control Kit (Affymetrix, cat. no. 900433).

2.5.2. Hybridization Controls

1. GeneChip Eukaryotic Hybridization Control Kit, Affymetrix, cat. no. 900454 (30 reactions) or cat. no. 900457 (150 reactions); contains control cRNA and control Oligo B2.
2. 3 n*M* Control Oligo B2 (Affymetrix, cat. no. 900301).

2.6. Preparation of the Hybridization Cocktail and Hybridization

Note that Affymetrix now offers a GeneChip Hybridization, Wash, and Stain Kit that provides all necessary reagents required to complete the hybridization, wash, and staining processes for GeneChip brand arrays in cartridge format. The kit includes pre-formulated solutions to process 30 arrays on the GeneChip Fluidics Station 400 or 450.

1. Nuclease-free water (Ambion, cat. no. 9930).
2. 50 mg/mL Bovine serum albumin (BSA) solution (Invitrogen Life Technologies, cat. no. 15561-020).
3. Herring sperm DNA (Promega, cat. no. D1811).
4. 5 *M* NaCl, RNase-free, DNase-free (Ambion, cat. no. 9760G).

5. MES hydrate SigmaUltra (Sigma-Aldrich, cat. no. M5287).
6. MES sodium salt (Sigma-Aldrich, cat. no. M5057).
7. 0.5 *M* EDTA, pH 8.0 (Ambion, cat. no. 9260G).
8. Dimethyl sulfoxide (DMSO; Sigma-Aldrich, cat. no. D5879).
9. 10% Surfact-Amps 20 (Tween-20; Pierce Chemical, cat. no. 28320).
10. 12X MES stock buffer: 1.22 *M* MES, 0.89 *M* [Na^+]. 35.2 g MES hydrate, 96.65 g MES sodium salt, nuclease-free water to 500 mL; pH should be 6.5 to 6.7. Sterile filter through 0.2-µm filter, and store protected from light at 4°C.
11. 2X MES hybridization buffer: 100 m*M* MES, 1 *M* [Na^+], 20 m*M* EDTA, 0.01% Tween-20. 8.3 mL 12X MES buffer, 17.7 mL 5 *M* NaCl, 4 mL 0.5 *M* EDTA pH 8.0, 0.1 mL 10% Tween-20, 19.9 mL nuclease-free water. Store at 4°C, and discard if yellow.
12. 1.7 mL SafeSeal tubes (RNase-free; Sorenson, cat. no. 11510).
13. 0.1 to 10 µL sterile, filter-free pipet tips with sharp tips, not beveled.
14. Lab tape or Tough Spots (Diversified Biotech, T-Spots).

2.7. Washing, Staining, and Scanning of Affymetrix GeneChip Probe Arrays

1. Water, molecular biology grade (BioWhittaker Molecular Applications/Cambrex, cat. no. 51200).
2. Distilled water (Invitrogen, cat. no. 15230-147).
3. 50 mg/mL BSA solution (Invitrogen, cat. no. 15561-020).
4. R-Phycoerythrin Streptavidin (Molecular Probes, cat. no. S-866).
5. 5 *M* NaCl, RNase-free, DNase-free (Ambion, cat. no. 9760G).
6. PBS, pH 7.2 (Invitrogen Life Technologies, cat. no. 20012-027).
7. 20X SSPE: 3 *M* NaCl, 0.2 *M* NaH_2PO_4, 0.02 *M* EDTA (BioWhittaker Molecular Applications/Cambrex, cat. no. 51214).
8. Goat IgG, reagent grade (Sigma-Aldrich, cat. no. I 5256); prepare 10 mg/mL stock by resuspending 50 mg in 5 mL of 150 m*M* NaCl. Aliquot and store at –20°C.
9. Antistreptavidin antibody (goat), biotinylated (Vector, cat. no. BA-0500). Prepare 0.5 mg/mL stock in sterile, nuclease-free water.
10. 10% Surfact-Amps 20 (Tween-20) (Pierce Chemical, cat. no. 28320).
11. 1.5-mL Sterile, RNase-free, microcentrifuge vials (USA Scientific, cat. no. 1415-2600).
12. Lab tape or Tough Spots (Diversified Biotech, T-Spots).
13. Wash Buffer A (nonstringent Wash Buffer): 300 mL of 20X SSPE, 1 mL of 10% Tween-20, 699 mL water. Filter through a 0.2-µm filter, and store protected from light at 4°C.
14. Wash Buffer B (stringent Wash Buffer): 83.3 mL of 12X MES stock buffer, 5.2 mL of 5 *M* NaCl, 1 mL of 10% Tween-20, 910.5 mL of water. Filter through a 0.2-µm filter, and store protected from light at 4°C.
15. 2X Stain Buffer: 41.7 mL of 12X MES stock buffer, 92.5 mL of 5 *M* NaCl, 2.5 mL of 10% Tween-20, 113.3 mL of water. Filter through a 0.2-µm filter, and store protected from light at 4°C.

3. Methods

3.1. Experimental Design

Microarray experimental design is a critical and mostly overlooked key element in a successful microarray experiment. The requirements for experimental design are no different for cardiovascular experiments than for any other field of application. The best designed microarray experiments begin with determining the objectives of the experiment and some well-defined goals. An experimental design that addresses a key hypothesis rather than being overly complex minimizes the arrays required and simplifies the data analysis. If several questions are to be addressed at once, the approach has to have enough statistical power, e.g., sufficient replicates, to answer all questions. This means that for each new variable added to a design, the required number of arrays is multiplied. Pilot microarray studies that focus on a single variable versus a control state are useful for first-time microarray users. They can help to identify problems related to the biological sample, the procedures, and the data analysis. Refining methods after a small-scale study is far cheaper and more effective than complex mathematical fixes after the fact. Pilot studies also provide a good estimate of the variance of gene expression, which is useful in determining how many replicates the experiment's key questions will require.

3.2. Heart Tissue Isolation from Mice or Rats

1. Anesthetize or sacrifice the animal according to the animal care and use guidelines of your institution and National Institute of Health standards.
2. Generously wet the chest of the animal with 70% alcohol, and wipe with sterile gauze to prevent the transfer of animal hair to the instruments and inside the body cavity.
3. Beginning 1 in. from the base of the sternum, open the body cavity up to the salivary glands.
4. Remove the sternum and move the lung to the side. Carefully lift up the heart with forceps and clip the dorsal aorta. Quickly remove the heart.
5. Adult heart tissue is rinsed in ice-cold PBS to remove blood, briefly blotted on a piece of gauze and immediately snap frozen in liquid nitrogen (*see* **Note 1**). Embryonic tissue is dissected in ice-cold PBS, and heart tissue is frozen in liquid nitrogen or on dry ice immediately after dissection.

3.3. RNA Extraction from Cardiac Tissues

Total RNA is isolated using the TRIZOL reagent, a monophasic solution of phenol and guanidine isothiocyanate, following the instructions of the manufacturer (*see* **Note 2**). Depending on the amount of starting material, the isolated RNA is then further purified by a phenol-chloroform extraction followed

by ethanol precipitation, or by binding to a silica gel-based membrane using RNeasy columns.

1. Prechill 100 mL PBS and a 75-mm tube with TRIZOL reagent on ice.
2. Weigh the tissue while taking care to keep it frozen (*see* **Note 3**). An adult mouse heart weighs about 100 to 170 mg. Homogenize tissue samples in 1 mL of TRIZOL reagent per 50 to 100 mg of tissue using a power homogenizer with a sawtooth generator probe (*see* **Note 4**). The sample volume should not exceed 10% of the volume of TRIZOL reagent used for homogenization.
3. Transfer samples into 75-mm tubes filled with the appropriate amount of room temperature (RT) TRIZOL reagent using a prechilled spatula (*see* **Note 5**). The minimum volume needed for these tubes is 1 mL.
4. Homogenize at full speed for 3 × 20 s while the sample is still frozen. Allow the sample to cool between homogenization steps.
5. Let the sample sit at room temperature for 5 min after homogenization to allow for complete dissociation of protein complexes.
6. Clean the generator probe in between samples by running the generator at full speed in the following solvents:

   30 s in 100% ethanol.
   30 s in chilled PBS 1.
   30 s in chilled PBS 2.
   10 s in chilled TRIZOL reagent.

   Dry well with Kimwipes
7. Add 0.2 mL of chloroform per 1 mL of TRIZOL reagent, and shake vigorously for 15 s.
8. Immediately transfer the contents into an RNase-free Eppendorf tube.
9. Incubate at RT for 3 min and centrifuge at 12,000*g* for 15 min at 4°C. Centrifugation separates the lower, red phenol-chloroform phase from the interphase and the clear aqueous upper phase, which contains the RNA.
10. Transfer the aqueous phase to a new Eppendorf tube.

    Important: For small starting amounts of tissue (e.g., cardiac needle biopsies) further purify the RNA by performing an additional phenol-chloroform extraction following **steps 11** to **16** in **Subheading 3.3.1.** below. For larger tissue starting amounts (e.g., using a whole adult mouse heart), use RNeasy columns and follow **steps 17** to **35** in **Subheading 3.3.2.**

3.3.1. Phenol-Chloroform Extraction

11. Add an equal amount (600 µL, if 1 mL of TRIZOL was used for the initial homogenization) of phenol-chloroform and vortex for 1 min. Centrifuge at top speed for 5 min at RT.
12. To remove phenol, transfer the aqueous phase to a new Eppendorf tube. Add an equal amount (600 µL) of isoamyl alcohol/chloroform and vortex for 1 min. Centrifuge at top speed for 5 min at RT.

13. To precipitate the RNA, transfer the aqueous phase to a new microcentrifuge tube and add 0.6 to 0.7 vol of isopropanol (approx 700 µL for 1 mL of TRIZOL in the initial homogenization). Mix by inverting 6 to 8 times and incubate at RT for 10 min.
14. Centrifuge at 12,000*g* for 10 min at 4°C, discard the supernatant, and wash the pellet with 1 mL of 75% ethanol in DEPC-treated water.
15. Centrifuge at 7500*g* for 5 min at 4°C, discard the supernatant, and allow the pellet to air-dry for 10 min at RT.
16. Dissolve the pellet in DEPC-treated water to a final concentration of 1 µg/µL. The expected yield of RNA is about 1 µg of RNA per mg of tissue.

3.3.2. Purification of RNA Using RNeasy Columns

17. The RNeasy spin columns are used to purify RNA after TRIZOL extraction (*see* **Subheading 3.3.**, **steps 1** to **10**) when the amount of starting material was at least 50 mg of tissue. Guanidine isothiocyanate (GITC)-containing lysis buffer and ethanol are added to the sample to create conditions that promote selective binding of RNA to the RNeasy membrane. Contaminants are washed away, and the purified RNA is then eluted from the membrane in DEPC-treated water.
18. Precipitate the RNA by adding 0.6 to 0.7 vol of isopropanol (approx 700 µL for 1 mL of TRIZOL in the initial homogenization) to the supernatant of **step 11** in **Subheading 3.3.1.** Mix by inverting 6 to 8 times and incubate at RT for 10 min.
19. Centrifuge at 12,000*g* for 10 min at 4°C, discard the supernatant, and wash the pellet with 1 mL of 75% ethanol in DEPC-treated water.
20. Centrifuge at 7500*g* for 5 min at 4°C, discard the supernatant, and allow the pellet to air-dry for 10 min at RT.
21. Estimate the amount of RNA in the pellet based on the assumption that 1 µg RNA can be extracted per mg tissue. The binding capacity of an RNeasy mini spin column is 100 µg of RNA. Therefore, dissolve up to 100 µg RNA in 100 µL RNase-free water. If the expected RNA yield is larger than 100 µg, use multiple RNeasy columns per sample.

The following steps are carried out at RT. Work quickly through the whole cleanup procedure to minimize the risk of RNA degradation.

22. Add 10 µL β-mercaptoethanol (β-ME) per 1 mL buffer RLT before use.

    Caution: β-ME is toxic; dispense in a fume hood and wear appropriate protective clothing.

    Buffer RLT is stable for 1 mo after addition of β-ME.
23. Add 350 µL Buffer RLT to the 100 µL RNA-containing solution, and mix thoroughly.
24. Add 250 µL ethanol (96–100%) to the diluted RNA, and mix thoroughly by pipeting.
25. Immediately apply the sample (700 µL) to an RNeasy mini column placed in a 2-mL collection tube (supplied with the kit).

26. Close the tube and centrifuge for 15 s at ≥8000*g* (≥10,000 rpm). Discard the flowthrough and collection tube.
27. Buffer RPE is supplied as a concentrate. Before using for the first time, add 4 vol of ethanol (96–100%), as indicated on the bottle, to obtain a working solution.
28. Transfer the RNeasy column into a new 2-mL collection tube. Pipet 500 µL Buffer RPE onto the RNeasy column. Close the tube and centrifuge for 15 s at ≥8000*g* (≥10,000 rpm) to wash the column. Discard the flowthrough, and reuse the collection tube.
29. Add another 500 µL Buffer RPE to the RNeasy column. Close the tube and centrifuge for 2 min at ≥8000*g* (≥10,000 rpm) to dry the RNeasy silica-gel membrane.
30. Discard the flowthrough, and reuse the collection tube. Centrifuge the column in the collection tube in a microcentrifuge at full speed for 1 min. It is important to dry the RNeasy silica-gel membrane since residual ethanol will interfere with downstream reactions.
31. After the centrifugation, place the RNeasy mini column in a new 1.5-mL collection tube (supplied with the kit).
32. Depending on the expected RNA yield, add 30 to 50 µL RNase-free water directly onto the RNeasy silica-gel membrane.
33. Close the tube and centrifuge for 1 min at ≥8000*g* (≥10,000 rpm) to elute the RNA.
34. Repeat the elution step (**step 33**) with another 30 to 50 µL RNase-free water, and elute into the same collection tube.
35. Determine the concentration, yield, and purity (*see* **Subheading 3.4.**) and freeze the samples at –80°C until the RNA is to be used for the labeling process (*see* **Note 6**).

3.4. RNA Quality Control and Quantitation

Impurities in RNA samples have an adverse effect on both the labeling efficiency and the stability of the fluorescent labels that are used. Therefore, the success of a microarray experiment largely depends on the quality of the prepared RNA. It is important to check the quality of the RNA after each thawing and before starting the labeling reaction, since impure RNA samples might degrade during the thawing process.

One measure of RNA purity is the ratio of absorbance readings at 260 and 280 nm. The A260:A280 ratio for RNA samples of acceptable purity should be between 1.8 and 2. RNA preparations should be free of contaminating proteins and organic solvents such as phenol or ethanol, and salts; they should also show no signs of degradation. A high-quality preparation of mammalian total RNA is characterized by two bright bands at approx 4.5 and 1.9 kb, representing 28S and 18S ribosomal RNA, and the absence of genomic DNA. A simple test for genomic DNA contamination is to use your RNA directly as a template

in a polymerase chain reaction (PCR) reaction with primers for any well-characterized gene, such as β-actin or GAPDH. The primer positions should not span a large intron and should be chosen to amplify a short fragment (<1 kb in length). RNA subjected to a reverse transcription reaction can be used as a positive control. If the PCR produces a visible band on an ethidium bromide-stained agarose gel, the RNA preparation contains genomic DNA. For a successful microarray experiment, the genomic DNA content in the RNA preparation should be less than 0.001% and should not produce a visible band after 35 PCR cycles.

The integrity of the RNA can be determined by running an aliquot of the RNA preparation on a 1% denaturing agarose gel, or, more quantitatively, using the Agilent 2100 Bioanalyzer and RNA 6000 Nano Labchip or RNA 6000 Pico Labchip. This method requires only nanogram amounts of RNA and is therefore especially suitable if the amount of RNA is limited. The protocol below provides instructions for determining the integrity and concentration of RNA samples on the Agilent 2100 Bioanalyzer using the RNA Nano Protocols. The manufacturer's detailed protocol is freely available through their website at www.chem.agilent.com.

3.4.1. Preparing the Reagents

1. Prepare the gel by placing 400 μL of RT RNA 6000 Nano gel matrix onto a spin filter column.
2. Place the spin filter in an Eppendorf tube and spin for 10 min at 1500*g*. The filtered gel should be stored at 4°C and used within a month after preparation.
3. Prepare the gel-dye mix by adding 2 μL of dye to 130 μL of filtered gel. All reagents should be at RT. Vortex and spin down at 13,000*g* for 10 min. To protect the gel-dye mix from light, cover the tube with foil. The gel-dye mix should be used within 1 wk from the date of preparation.
4. Prepare the RNA 6000 ladder by denaturing an aliquot of the ladder for 2 min at 70°C. Snap-cool on ice. For each chip, 1 μL of the ladder is needed.

3.4.2. Preparing the Bioanalyzer

The electrodes should be decontaminated before each run.

1. Place a washing chip (electrode cleaner) filled with 350 μL RNaseZap in the Bioanalyzer and close the lid for 1 min. Remove the washing chip from the Bioanalyzer.
2. Place another washing chip filled with 350 μL DEPC-water in the Bioanalyzer, and close the lid for 30 s.
3. Remove the washing chip and leave the lid open for 10 s to allow the electrodes to dry.

3.4.3. Loading the Chip

Remove the chip from packaging and inspect the under-side wafer of the chip for any defects.

1. Place the chip into the priming station. Pull out the syringe to the 1-mL mark.
2. Pipet 9 µL of gel-dye mix into the well marked with a back-circled "G." Always place the pipet tip into the center and bottom of each well when dispensing.
3. Snap the priming station lid closed.
4. Push the plunger of the syringe down to 0.2 mL, so that it fits snugly under the silver stopper. Press down slowly and steadily on the plunger when priming.
5. Wait exactly 30 s.
6. Squeeze the silver stopper to release the plunger. It should come up to at least 0.7 mL within 1 to 2 s. If it does not prime well, check that the gasket is clean. Change the gasket if the problem persists.
7. Pull the plunger up the rest of the way to 1 mL. Be careful to avoid a negative pressure vacuum, which can lead to bubbles.
8. Lift the priming station lid. Hold the chip in place with one hand and lift the lid with the other to prevent dislodging of the chip from the priming station.
9. Pipette 9 µL of gel-dye mix into each one of the other two wells labeled "G."
10. Pipette 5 µL of the RNA 6000 Nano Marker into the well marked with the ladder symbol and into each of the 12 sample wells.
11. Do not leave any wells empty. Add 6 µL of the RNA Nano Marker to each unused well (*see* **Note 7**).

3.4.4. Starting the Chip Run

1. The quantitative measurement range for total RNA concentrations is 25 to 500 ng/µL, and that for mRNA concentrations is 25 to 250 ng/µL. For qualitative measurements to check for RNA integrity only, the range is 5 to 500 ng/µL.
2. To minimize secondary structure, heat denature (70°C, 2 min) the samples before loading on the chip.
3. Add 1 µL of sample into each well. Empty wells should be filled with 6 µL water or the Nano marker. Accurate pipeting is very important. Use properly calibrated pipets. You may pipet up and down gently to mix samples in the wells.
4. Pipet 1 µL of the ladder into the ladder well.
5. Vortex the chip for 1 min at about 512*g*. Use a piece of tape or an elastic band to secure the chip in the adapter. This will prevent the chip from falling out even if it has been inserted into the vortexer adapter in the wrong orientation.
6. Use the chip within 5 min of preparation to prevent evaporation. Cover the chip if it will be left standing for any length of time
7. Launch the 2100 Expert software.
8. Place the chip into the Bioanalyzer and close the lid. A chip symbol will appear on the screen.
9. Click the *Electrophoresis* button, click *Assay*, click *RNA*, and select *Eukaryote Total RNA Nano* from the menu.

10. If necessary, edit the number of samples using the menu on the lower left.
11. Press the *Start* button in the upper right to start the chip run.
12. Edit the sample names by clicking the blue *Data File* link or click the *Data and Assay* context.
13. Select the *Chip Summary* tab, and enter the sample names into the table.
14. Press *Apply* in the lower right hand corner when you are finished.
15. Click the *Instrument* context to review the live trace of your samples.
16. The approximate run time per sample is 90 s. A full chip run takes about 20 min.

3.4.5. Cleaning and Maintenance

1. Remove the chip from the Bioanalyzer immediately after the run is completed. Do not let the sample dry on the electrodes.
2. After each run, follow the cleaning **steps 1** to **3** in **Subheading 3.4.2.**, "Preparing the Bioanalyzer."
3. The focus lens should be cleaned monthly with isopropanol.
4. The electrodes should be cleaned at least quarterly by sonication.

3.4.6. Interpreting RNA Nano Results

1. To check the results of the run, select the *Gel* or *Electropherogram* tab in the *Data and Assay* context.
2. Select the ladder. The electropherogram of the ladder should show six RNA peaks (depending on the peak finding settings, the last peak might not be detected) and one marker peak. All seven peaks should be well resolved.
3. To review the data for a specific sample, select the sample name and highlight the *Results* subtab. Major features of an electropherogram for a successful eukaryotic total RNA run are sharp and well-resolved 18S and 28S ribosomal RNA peaks. There should be no or only a small peak for the 5S ribosomal RNA. The baseline should be flat. Smaller additional peaks and a hump in the baseline are indicative of degradation (*see* **Note 8**).
4. Microarray gene expression experiments should not be performed with poor-quality total RNA samples.

3.5. RNA Amplification and Labeling

Total RNA or mRNA (*see* **Note 9**) is first reverse-transcribed using a T7 promoter-containing oligo(dT) primer in the first-strand cDNA synthesis reaction. Following RNase H-mediated second-strand cDNA synthesis, the double-stranded cDNA is purified and serves as a template in the subsequent in vitro transcription (IVT) reaction. The IVT reaction is carried out in the presence of T7 RNA polymerase and a biotinylated nucleotide analog/ribonucleotide mix for complementary RNA (cRNA) amplification and biotin labeling.

3.5.1. First-Strand cDNA Synthesis

1. Perform the steps outlined in **Subheading 3.6.1.** if poly-A RNA spike-in controls are being used.
2. Briefly spin down all tubes in the kit before using the reagents.
3. For the first-strand cDNA synthesis, program a thermal cycler with heated lid as follows: 70°C for 10 min, 4°C hold, 42°C for 2 min, 42°C for 1 h, 4°C hold. The 4°C holds are for reagent addition steps.
4. Pipet the total RNA sample (5 µg) into an RNase-free thin-wall 0.2-mL PCR tube. If the reaction includes the poly-A RNA controls, the volume of the total RNA should comprise no more than 8 µL (*see* **Note 10**).
5. Add 2 µL of the appropriately diluted poly-A RNA controls (*see* **Subheading 3.6.1.**), 2 µL of 50 µ*M* T7-Oligo(dT) Primer (included), and RNase-free water to a final volume of 12 µL.
6. Gently flick the tube a few times to mix, and then centrifuge briefly (approx 5 s) to collect the reaction at the bottom of the tube.
7. Perform the first incubation step in the thermal cycler, e.g., incubate for 10 min at 70°C. Cool the sample after the incubation at 4°C at least 2 min.
8. Prepare sufficient First-Strand Mastermix for all RNA samples: 4 µL 5X First Strand Reaction Mix, 2 µL 0.1 *M* dithiothreitol (DTT), 1 µL 10 m*M* dNTP.
9. Add 7 µL First-Strand Mastermix to each RNA/T7-Oligo(dT) Primer mix for a final volume of 19 µL while the samples are on the 4°C hold step. Mix well by flicking the tubes a few times, and centrifuge briefly to collect the reaction at the bottom of the tube.
10. Incubate for 2 min at 42°C.
11. Add 1 µL SuperScript II reverse transcriptase to each RNA sample for a final volume of 20 µL.
12. Incubate at 42°C for 1 h.
13. Place the samples on ice for at least 2 min, and immediately proceed with the second-strand synthesis.

3.5.2. Second-Strand cDNA Synthesis

1. For the second-strand cDNA synthesis, program a thermal cycler as follows: 16°C for 2 h, 4°C hold, 16°C for 5 min, 4°C hold.
2. Prepare a second-strand mastermix for all the samples. The following recipe is for a single reaction: 91 µL RNase-free water, 30 µL 5X second-strand reaction mix, 30 µL 10 m*M* dNTP, 1 µL *E. coli* DNA ligase, 4 µL *E. coli* DNA polymerase I, 1 µL RNaseH. The total volume is 130 µL.

   Mix well and centrifuge briefly to collect the mix at the bottom of the tube.
3. Add 130 µL of the second-strand mastermix to each first-strand synthesis reaction from **Subheading 3.5.1.** Mix well and briefly centrifuge.
4. Return the tubes to the thermal cycler and incubate for 2 h at 16°C.
5. Add 2 µL of T4 DNA polymerase to each sample and continue incubation at 16°C for another 5 min.

6. Add 10 µL of 0.5 *M* EDTA and place the reactions on ice. The total volume is now 162 µL.
7. Do not leave the samples on ice for a prolonged period. Samples may be stored at –20°C for later use. However, it is recommended to proceed directly to **Subheading 3.5.3.** just below to avoid unnecessary freeze/thaw cycles that may impact the integrity of the cDNA.

3.5.3. Cleanup of Double-Stranded cDNA

The components needed for cleaning up the cDNA are supplied with the GeneChip Sample Cleanup Module. Add 24 mL of 96 to 100% ethanol to the cDNA Wash Buffer prior to the first use.

1. Add 600 µL of cDNA Binding Buffer to the second strand synthesis reaction and mix by vortexing. Check that the color of the mixture is yellow, similar to the cDNA Binding Buffer. Add 10 µL of 3 *M* NaOAc, pH 5.0, if the color is orange or violet.
2. Apply 500 µL of the sample to the cDNA Cleanup Spin Column sitting in a 2-mL Collection Tube (supplied), and centrifuge for 1 min at ≥8000*g* (≥10,000 rpm).
3. Discard the flowthrough, reload the spin column with the remaining mixture (262 µL), and centrifuge as above.
4. Discard the flowthrough, and transfer the spin column into a new 2-mL Collection Tube (supplied).
5. Pipet 750 µL of the cDNA Wash Buffer onto the spin column. Centrifuge for 1 min at ≥8000*g* (≥10,000 rpm). Discard flowthrough.
6. Open the cap of the spin column and centrifuge for 5 min at maximum speed (≤25,000*g*) to dry the membrane completely. Discard flowthrough and Collection Tube.
7. Transfer the spin column into a 1.5-mL Collection Tube, and pipet 14 µL of cDNA Elution Buffer directly onto the spin column membrane.
8. Incubate for 1 min at room temperature and centrifuge for 1 min at maximum speed (≤25,000*g*) to elute. The average volume of eluate is 12 µL (*see* **Note 11**).
9. After cleanup, proceed to **Subheading 3.5.4.** just below.

3.5.4. Synthesis of Biotin-Labeled cRNA by In Vitro Transcription

The GeneChip IVT Labeling Kit is used for this step.

1. Prepare a mastermix for all the samples. The following recipe is for a single reaction: 12 µL template cDNA from **Subheading 3.5.3.** (*see* **Note 12**), 4 µL 10X IVT Labeling Buffer, 12 µL IVT Labeling NTP Mix, 4 µL IVT Labeling Enzyme Mix, 8 µL RNase-free water. The total volume is 40 µL.
2. Mix well and incubate at 37°C for 16 h in a thermal cycler with heated lid to avoid condensation.
3. Labeled cRNA can be stored at –20°C for short-term storage or at –70°C for longer periods until further purified. However, it is recommended to proceed

directly to **Subheading 3.5.5.** just below after the 16-h incubation to avoid freeze/thaw cycles.

3.5.5. Cleanup and Quantification of Biotin-Labeled cRNA

To determine the concentration of the cRNA accurately, it is necessary to remove unincorporated NTPs. The Sample Cleanup Module is used for this purpose (*see* **Note 13**). Before using the IVT cRNA Wash Buffer for the first time, add 20 mL of ethanol (96–100%), as indicated on the bottle, to obtain a working solution.

1. Add 60 μL of RNase-free water to the IVT reaction, and mix by vortexing. The total volume is now 100 μL.
2. Add 350 μL IVT cRNA Binding Buffer to the sample and mix by vortexing.
3. Add 250 μL 96 to 100% ethanol to the lysate, and mix well by pipetting. Do not centrifuge.
4. Apply the sample (700 μL) to the IVT cRNA Cleanup Spin Column sitting in a 2-mL Collection Tube. Centrifuge for 15 s at ≥8000*g* (≥10,000 rpm).
5. Discard flowthrough and collection tube. Transfer the spin column into a new 2-mL Collection Tube (supplied).
6. Pipet 500 μL IVT cRNA Wash Buffer onto the spin column. Centrifuge for 15 s at ≥8000*g* (≥10,000 rpm) to wash. Discard flowthrough.
7. Pipet 500 μL 80% (v/v) ethanol onto the spin column and centrifuge for 15 s at ≥8000*g* (≥10,000 rpm). Discard flowthrough.
8. Open the cap of the spin column and centrifuge for 5 min at maximum speed (≤25,000*g*) to dry the membrane completely. Discard flowthrough and Collection Tube.
9. To elute the cRNA, transfer the spin column into a new 1.5-mL Collection Tube (supplied), and pipet 11 μL of RNase-free water directly onto the spin column membrane. Centrifuge for 1 min at maximum speed (≤25,000*g*).
10. Determine the cRNA yield by measuring absorbance at 260 nm, and determine the purity by measuring the 260:280 nm ratio using a spectrophotometer or the Nanodrop ND-1000.
11. To estimate the size distribution, run a 1 ng aliquot of the purified cRNA on the Agilent Bioanalyzer using the RNA labchip (*see* **Subheading 3.4.**) (*see* **Note 14**).

3.5.6. Fragmentation of the Biotin-Labeled cRNA

To achieve optimal binding sensitivity on Affymetrix GeneChip probe arrays, the biotin-labeled cRNA is fragmented by metal-induced hydrolysis into 35 to 200 base fragments (*see* **Note 15**). The Sample Cleanup Module is used for this step. It is usually best to perform the fragmentation on the same day as hybridization whenever possible to minimize freeze/thawing effects on the cRNA. The degree of fragmentation can vary substantially. Therefore, exact timing and accurate pipeting techniques are crucial.

1. For the fragmentation, program a thermal cycler with heated lid as follows: 94°C for 35 min, 4°C hold.
2. Prepare the following fragmentation mix: 20 µg cRNA (adjust the concentration from **Subheading 3.5.5.** to reflect carryover of the unlabeled total RNA input; the volume of cRNA should not exceed 21 µL), 8 µL 5X Fragmentation Buffer. Add RNase-free water to 40 µL final volume.
3. Incubate at 94°C for exactly 35 min. Put the sample on ice immediately after incubation.
4. Check the cRNA by the Agilent Bioanalyzer using the mRNA Smear Nano Assay (*see* **Subheading 3.4.**). Run 100 ng of unfragmented cRNA side by side with 1 µL fragmented undiluted cRNA directly from the fragmentation reaction. The assay should reveal long unfragmented cRNAs greater than 400 bp, whereas the average size of the fragmented cRNA should be around 100 nucleotides.
5. The remaining fragmented cRNA should immediately be used to assemble the hybridization cocktail. Do not leave reactions at 4°C for long periods. However, it is possible to store the assembled hybridization mix at –80°C until the day of hybridization.

3.6. Controls

Affymetrix provides two kinds of controls. The poly-A RNA controls correspond to probe sets on the GeneChip arrays for *Bacillus subtilis* genes (*lys*, *phe*, *thr*, and *dap*) that are not present in eukaryotic samples. These spike-in controls are added to the isolated RNA samples and then amplified and labeled together with the samples. Comparing the hybridization intensities of these controls on all GeneChip arrays in an experiment allows one to monitor the entire labeling process. However, the hybridization signals from poly-A spike-in controls do not allow one to assess the initial quantity and quality of the isolated RNA samples.

The other type is hybridization controls (*bioB*, *bioC*, *bioD*, *cre*) that are added to the hybridization mix. These controls only allow an evaluation of the efficiency of hybridization.

3.6.1. Poly-A RNA Spike-In Controls

Use the Eukaryotic Poly-A RNA Control Kit for this step. The in vitro synthesized transcripts in the kit are already premixed at different concentrations. The concentrated Poly-A Control Stock is diluted using Poly-A Control Dilution Buffer (provided in the kit) and spiked directly into RNA samples to achieve the following final concentrations, given as the ratio of copy numbers: *lys*, 1:100,000; *phe*, 1:50,000; *thr*, 1:25,000; *dap*, 1:6,667.

1. For a starting amount of 5 µg total RNA, the Poly-A Control Stock is diluted 1:10,000 in three serial dilutions of 1:20, 1:50, and 1:10.

2. Then 2 µL of the diluted stock are added to the 5 µg total RNA prior to labeling (*see* **Note 16**).

3.6.2. Hybridization Controls

The Eukaryotic Hybridization Control Kit contains a mixture of biotin-labeled cRNA transcripts of *bioB*, *bioC*, *bioD*, and *cre*, prepared in staggered concentrations (1.5, 5, 25, and 100 p*M*, respectively). *bioB*, *bioC*, and *bioD* represent genes in the biotin synthesis pathway of *E. coli*, and *cre* is a recombinase gene of bacteriophage P1. The Eukaryotic Hybridization Controls are spiked into the hybridization cocktail as described in **Subheading 3.7.** just below and are thus used to evaluate sample hybridization efficiency of the probe array.

3.7. Preparation of the Hybridization Cocktail and Hybridization

It is usually best to make all cocktails for a set of arrays fresh to minimize freeze/thawing effects on cRNA and at the same time minimize day-to-day variability.

The final concentration of the cRNA in the hybridization cocktail must be no less than 0.05 µg/µL. If the recommended fragmentation reaction (20 µg of cRNA fragmented in 40 µL) was followed, the sample meets this requirement.

3.7.1. Equilibration of Reagents

1. Equilibrate arrays to room temperature approx 15–20 min before use. Failure to do so may cause the microarray septa to crack and damage the microarray irreversibly.
2. Preheat the hybridization oven to 45°C (takes about 5 min).
3. Equilibrate the Hybridization Buffer to room temperature for 5–10 min before use.
4. Heat the 20X Eukaryotic Hybridization Controls and OligoB2 to 65°C in 1.7-mL microcentrifuge tubes for 5–10 min to ensure that the cRNA is completely resuspended. Briefly vortex to mix and spin down to collect on the bottom of the tube.
5. Thaw the Herring Sperm DNA and BSA at 37°C for 5–10 min. Briefly vortex to mix and spin down to collect on the bottom of the tube.
6. Keep the fragmented biotin-labeled cRNA from **Subheading 3.5.6.** on ice until ready to use.

3.7.2. Preparing the Affymetrix GeneChip Array

1. Remove the array from its packaging and inspect it for manufacturing defects (i.e., scratches on glass or wafer). Record the array type, lot number, and expiration date.
2. Wet the array by filling it through one of the septa with 200 µL of 1X Hybridization Buffer. It is necessary to use two pipet tips when filling the probe array

Table 1
Hybridization Cocktail for Standard Affymetrix GeneChip Probe Arrays

Component	Volume (µL)	Final concentration
Control oligonucleotide B2 (3 n*M*)	5	0.05 n*M*
20X Eukaryotic hybridization controls	15	1X (1.5, 5, 25, and 100 p*M*, respectively)
Herring Sperm DNA (10 mg/mL)	3	0.1 mg/mL
BSA (50 mg/mL)	3	0.5 mg/mL
2X Hybridization buffer	150	1X
100% Dimethyl sulfoxide (DMSO)	30	10%
Water	56	
Fragmented biotin-labeled cRNA (20 µg)	38	
Final volume	300	

cartridge: one for filling and the second to allow venting of air from the hybridization chamber. Use 100 µL for Test3 arrays and 300 µL for a standard GeneChip format.

3. Insert the array into the GeneChip cartridge carrier. Prepare a balanced carrier.
4. Insert the carriers into the preheated 45°C oven. Make sure that the carriers are properly secured.
5. Set the rotation speed to 60 rpm and prehybridize the microarrays for a minimum of 10 min but no longer than 25 min. Do not allow the microarrays to dry.

3.7.3. Preparation of the Hybridization Cocktail and Hybridization

1. Prepare the hybridization cocktail for a single probe array hybridization as shown in **Table 1**. The recipe takes into account that it is necessary to make extra hybridization cocktail owing to a small loss in volume during each hybridization. Scale up the volumes for hybridization to multiple arrays. Add the cRNA last. If an error is made, you will not have to redo the fragmentation step.
2. Heat the hybridization cocktail to 99°C for 5 min prior to use.
3. Transfer the hybridization cocktail to a 45°C heat block for 5 min to equilibrate the cocktail to the microarray temperature.
4. Spin the hybridization cocktail at maximum speed in a microcentrifuge for 5 min to remove any insoluble material from the hybridization mixture.
5. Remove the buffer solution from the probe array cartridge and fill with 200 µL of the clarified hybridization cocktail, avoiding any insoluble matter at the bottom of the tube. Upon filling, a small meniscus should be visible through the glass window, with few to no bubbles.
6. Seal both septa with lab tape or Tough-Spots.
7. Place the probe array into the prewarmed Hybridization Oven set to 45°C, and rotate at 60 rpm.
8. Hybridize for 16 h. Do not allow the microarrays to hybridize longer or dry.

3.8. Washing, Staining, and Scanning of Affymetrix GeneChip Probe Arrays

Washing and staining of Affymetrix GeneChip probe arrays is an automated process and requires the Affymetrix Fluidics station. The staining protocol includes a signal amplification step that employs a fluorescent molecule (streptavidin phycoerythrin) that binds to biotin, an antistreptavidin antibody (goat), and a biotinylated goat IgG antibody. A series of washes and stains with these reagents provides an amplified flour that emits light when the probe array is scanned with a confocal laser. Both the Fluidics Station and the scanner are controlled by the Affymetrix Microarray Suite (MAS) (older version) or the Affymetrix GeneChip® Operating Software (GCOS) (new version). GCOS also acquires data, manages sample and experimental information, and performs some gene expression data analysis.

The following protocols outline the general procedure for washing, staining, and scanning Affymetrix GeneChip arrays. Since there are several versions of the Fluidics Station (400 and 450/250), GeneArray and GeneChip scanner (2500 and 3000), and operating software (MAS and GCOS) in use, please refer to the specific Affymetrix User's Guides ***(2)*** for details on how to use the instrumentation and software.

3.8.1. Affymetrix GeneChip Probe Array Washing and Staining

Follow **steps 1** to **4** before taking the probe arrays out of the hybridization oven.

1. Launch the operating software and enter the experiment information according to the appropriate GCOS or Microarray Suite user's guides.
2. Set up and prime the Fluidics Station according to the Affymetrix User Manual for *Eukaryotic Sample and Array Processing* or the User's Guides of the instrument.
3. Prepare the streptavidin phycoerythrin (SAPE) stain solution (**Table 2**). The table lists the components needed for one to five probe arrays, adding a small volume for pipeting losses. Always prepare the stain solution fresh on the day of use, and prepare enough solution for all arrays to be processed on that day. SAPE should be stored in the dark at 4°C. Mix well and divide into two aliquots of 600 μL each per sample to be used for stains 1 and 3, respectively.
4. Prepare the antibody solution (**Table 3**). The table lists the components needed for one to five probe arrays, adding a small volume for pipeting losses. Mix well and use in aliquots of 600 μL to be used for stain 2. Prepare sufficient solution for all arrays to be processed on a given day.
5. After hybridization, recover the hybridization cocktail from the array and place it in the original tube. Store at –20°C for short-term storage or –80°C for long-term storage.
6. Fill the probe array completely with nonstringent Wash Buffer A (*see* **Note 17**).

Table 2
Streptavidin Phycoerythrin (SAPE) Stain Solution

Component	Final concentration	1	2.2	3.2	4.2	5.2
Water	—	540	1188	1728	2268	2808
2X Stain Buffer	1X	600	1320	1920	2520	3120
BSA, 50 mg/mL	2 mg/mL	48	105.6	153.6	201.6	249.6
SAPE, 1 mg/mL	10 µg/mL	12	26.4	38.4	50.4	62.4
Total (µL)		1200	2640	3840	5040	6240

Table 3
Antibody Solution

Component	Final concentration	1	2.2	3.2	4.2	5.2
Water	—	266.4	586.08	852.48	1118.88	1385.28
2X Stain Buffer	1X	300	660	960	1260	1560
BSA, 50 mg/mL	2 mg/mL	24	52.8	76.8	100.8	124.8
Goat IgG, 10 mg/mL	0.1 mg/mL	6	13.2	19.2	25.2	31.2
Biotinylated antibody, 0.5 mg/mL	3 µg/mL	3.6	7.92	11.52	15.12	18.72
Total (µL)		600	1320	1920	2520	3120

7. Start the Fluidics Station wash and stain run choosing the Fluidics scripts appropriate for your probe array. Insert the probe arrays into the designated module of the Fluidics Station; load the vial with the SAPE stain solution (Fluidics Station 400) or the two vials with the SAPE stain solution and the vial with the antibody solution (Fluidics Station 450/250) according to the Affymetrix User Manual for *Eukaryotic Sample and Array Processing*.
8. Proceed with the run and follow the prompts according to the protocol in the Affymetrix User Manual for *Eukaryotic Sample and Array Processing* for your instrument.
9. At the end of the run, remove the probe arrays from the Fluidics Station, and proceed with the probe array scan (*see* **Subheading 3.8.2.**).
10. If you do not scan the arrays right away, keep the probe arrays at 4°C and in the dark until ready for scanning. However, do not store arrays for a prolonged period, since this might lead to variability in the resulting data.
11. If there are no more samples to run for the day, shut down the Fluidics station following the procedure outlined in the Affymetrix User Manual for *Eukaryotic Sample and Array Processing*.

3.8.2. Scanning of Affymetrix GeneChip Probe Arrays

The scanner is also controlled by the Affymetrix Microarray Suite or GCOS.

1. Turn on the scanner at least 10 min before use to make sure that the laser is warmed up.
2. If the probe array was stored at 4°C, let it warm to room temperature before scanning.
3. Check that there are no bubbles visible in the glass window of the probe array. If bubbles are present, carefully remove the buffer and manually refill the array with nonstringent Wash Buffer A using a micropipet (*see* **Subheading 3.7.2.**, **step 2**).
4. Remove any dirt or dust from the glass surface of the probe array using a non-abrasive tissue (do not use alcohol or Kimwipe tissues). Apply Tough Spots to each of the two septa to prevent any leaking of fluids during the scanning. Ensure that the tape or the Tough Spots remain completely flat.
5. Follow the Microarray Suite or GCOS or the appropriate scanner's user's manual for inserting the probe arrays into the scanner and starting the scanning process.
6. After scanning, check the image for the presence of image artifacts (i.e., high/low intensity spots, scratches, high regional or overall background, and so on) on the array. Right-click the mouse while the cursor is placed over the image, choose *Image Settings*, and uncheck *AutoScale*. View the image at a setting of 0 to 1000 to detect areas with high background.
7. Save the Affymetrix GeneChip probe arrays after scanning until you have checked the image and made sure that there are no manufacturing defects. Store the array wrapped in aluminum foil at 4°C. If there are defects, consult with your local Affymetrix representative for a possible replacement.

3.8.3. File Types

The Affymetrix operating software (MAS or GCOS) generates a number of files during processing of an Affymetrix GeneChip experiment.

The **Experiment Information File** (*.exp) contains information about the experiment name, sample, and probe array type. This file is not used for analysis but may be required to open other files for the designated probe array.

The **Data File** (*.dat) is the image of the scanned probe array.

The **Cell Intensity File** (*.cel) is derived from a *.dat file and is automatically created when a *.dat file is opened. It contains a single-intensity value for each probe cell delineated by the grid (calculated by the Cell Analysis algorithm).

The **Chip File** (*.chp) is the output file generated from the analysis of a probe array. It contains qualitative and quantitative analysis for every probe set. The file can be converted to a tab-delimited text file for easy uploading into data analysis software.

The **Report File** (*.rpt) is a text file summarizing data quality information for a single experiment. The report is generated from the analysis output file (*.chp).

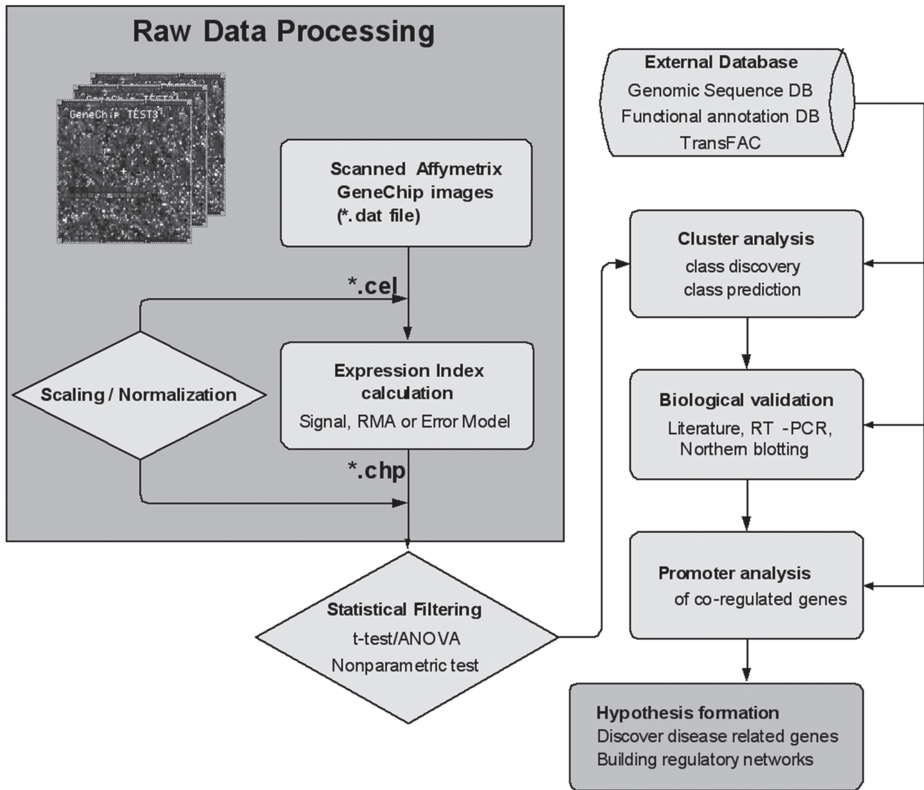

Fig. 1. Analysis steps involved in the first-order data analysis.

3.9. First-Order Data Analysis

Before the data can be used for statistical analysis, the raw values are scaled and normalized, and a gene expression index is calculated (**Fig. 1**).

Each transcript is measured by 11 probe pairs, which consist of a perfect match 25mer oligonucleotide (PM) and a 25mer mismatch oligonucleotide (MM) that contains a single base pair mismatch in the central position. The PM/MM design allows identification and subtraction of nonspecific hybridization and background signals, at the cost of introducing more variability in the gene expression indexes. Affymetrix GCOS uses a One-Step Tukey's Biweight Estimate to calculate the Signal, a quantitative metric that represents the relative level of expression of a transcript. It uses the MM intensity to estimate stray signal, which is subtracted from the PM. Further details about the analysis can be found in the Affymetrix User's Guides for MAS and GCOS, and the manual for *Data Analysis Fundamentals*.

Alternative approaches that use the PM intensity and other adjustment for background have been developed, such as Robust Multichip Average (RMA) ***(3)***, and Model-Based Expression Index (MBEI) implemented in dChip ***(4)***. It has been shown that the choice of the PM adjustment method can strongly influence the accuracy of the results ***(5)***. However, no method seems to be superior to the others in all aspects.

The Affymetrix software also calculates a Detection *p*-value, which is evaluated against user-definable cutoffs to determine the Detection call. The Detection call indicates whether a transcript is reliably detected (Present) or not detected (Absent).

In GCOS, the average Signal intensity of the array is set to a default Target Signal of 500. The key assumption of this global scaling strategy is that there are few changes in gene expression among all the genes on the arrays that are being analyzed. Although this is a widely used approach, it might not be suitable for all experiments.

3.10. Quality Control Criteria

Generating high-quality microarray data requires vigorous quality control measures at each individual step of the process, starting with the experimental design of the study, the generation of samples, extraction of RNA, labeling of the probe, and microarray hybridization.

RNA quality control measures are outlined in **Subheading 3.4.** Briefly, RNA purity and yield are determined by optical density (OD) measurements at wavelengths of 260 and 280 nm. The OD 260:280 ratio should lie between 1.7 and 2.0. Otherwise, the RNA should be repurified. Further evaluation of the RNA quality is done using the Agilent Bioanalyzer and Lab-on-a-Chip. Electropherograms are created that detect degradation and measure the ribosomal 5S, 18S, and 28S bands. Ideally, the ratio of 28S:18S bands should be close to 2, but samples that show clear 18S and 28S peaks are acceptable. An additional index for measuring RNA quality is the yield (mass conversion rate) and average size of the biotin-labeled cRNA after RNA labeling.

3.10.1. Array Hybridization Quality Control

1. A general *visual inspection* of the entire GeneChip probe array should be performed after scanning, as outlined in **Subheading 3.8.2.** Probe arrays that show white speckling, holes, smudges, areas of saturation, or uneven hybridization should be repeated, or the affected probe sets should be manually masked. Alternatively, the image can be evaluated and outliers detected using dCHIP software ***(4)***.
2. The boundaries of the probe area (viewed upon opening the *.dat/*.cel file) are easily identified by the hybridization of the *B2 Oligo*, which is spiked into each hybridization cocktail as part of the Eukaryotic Hybridization Control kit.

Hybridization of B2 is highlighted on the image by the alternating pattern of intensities on the border, the checkerboard pattern at each corner, and the array name, located in the upper left or upper middle of the array. B2 Oligo serves as a positive hybridization control and is used by the software to place a grid over the image. Some variation in B2 hybridization intensities across the array is normal. However, if the B2 intensities at the checkerboard corners are either too low or high, or are skewed owing to image artifacts, the grid will not align automatically. The user must align the grid manually using the mouse to click and drag each grid corner to its appropriate checkerboard corner.

3. The *Noise value (Raw Q)* can be found either in the Analysis Info tab of the Data Analysis (*.chp) file, or in the Expression Report (*.rpt) file. Noise (Raw Q) is a measure of the pixel-to-pixel variation of probe cells on a GeneChip array. RawQ values should remain consistent across the probe arrays of an experiment.
4. *Scaling factors (SF)* can be found in the Analysis Info tab of the .chp file output and in the Expression Report (.rpt) file. SF values should remain consistent across the experiment. The scaling factor for each given experiment should be within a two- to threefold range.
5. The number of probe sets called "*Present*" relative to the total number of probe sets on the array is displayed as a percentage in the Expression Report (.rpt) file. Percent Present (%P) values depend on multiple factors including cell/tissue type, biological or environmental stimuli, probe array type, and overall quality of RNA. Replicate samples should have similar %P values. Extremely low %P values are a possible indication of poor sample quality.
6. Most GeneChip expression arrays use β-*actin and GAPDH* as internal control genes. Transcripts of these genes are represented by 3' and 5' probe sets that are used to assess RNA sample and assay quality. The Signal values of the 3' probe sets are compared with the Signal values of the corresponding 5' probe sets. The ratio of the 3' probe set to the 5' probe set should be not more than 2 for the one-cycle assay. Additional rounds of amplification increase this number. A high 3' to 5' ratio may indicate degraded RNA or inefficient transcription of double-stranded cDNA. The 3' to 5' ratios for internal controls are displayed in the Expression Report (.rpt) file.
7. The *Eukaryotic Hybridization Controls bioB*, *bioC*, *bioD*, and *cre* are spiked in staggered concentrations (1.5, 5, 25, and 100 p*M* final concentrations for *bioB*, *bioC*, *bioD*, and *cre*, respectively) into the hybridization cocktail (*see* **Subheading 3.6.2.**). They should maintain a maximum 1:2 ratio of signal intensities of the 5' and 3' probe sets. *bioB* is at the level of assay sensitivity and should be called "Present" at least 50% of the time. *bioC*, *bioD*, and *cre* should always be called "Present" with increasing Signal values, reflecting their relative concentrations.
8. The *PolyA spike-in controls* (*see* **Subheading 3.6.1.**) should be called "Present" with increasing Signal values in the order of *lys*, *phe*, *thr*, and *dap*.

3.10.2. Graphical Quality Control

A number of graphical tools are available that allow one to visualize the quality of microarray data from several perspectives. These diagnostic plots include:

- A color image plot of the entire array to evaluate expression across the probe array's special layout.
- M versus A plots of two probe arrays display the log intensity ratio M versus the mean log intensity A. MvA plots make it easy to identify intensity-dependent biases in the data (i.e., "banana shape").
- Intensity boxplots are another good visualization tool for comparing the overall intensities of all probes across the arrays in an experiment.
- Intensity histograms are a good visualization tool for the distribution of intensities and for identifying saturation, which is revealed as an overrepresentation of high-intensity values.
- RNA degradation plots show expression as a function of the 3' to 5' position of probes from the *.cel file. The slope of these lines indicates potential RNA degradation in the samples.
- Principal components plots can help to evaluate inner group and between-group variability.

4. Notes

1. It is important that the tissue be snap-frozen immediately after tissue harvest. If a more time-intensive tissue dissection is required, adult heart tissue should be dissected and stored in RNA-stabilizing agents such as RNAlater (Qiagen) to preserve the RNA. RNAlater is not recommended for storage of embryonic tissue for subsequent dissection because it makes the tissue brittle.
2. For adult heart tissue, we found that phenol-based RNA isolation methods give higher yields than solid phase/glass filter-based methodologies alone. An additional purification step is necessary to yield RNA of high purity suited for microarray experiments. If sufficient starting tissue material is available, the use of a combination of phenol-based and solid phase/glass filter-based methods yields about 1 μg of high-purity RNA per mg of tissue.
3. If you are working with very small amounts of tissue, such as needle biopsy material, skip the weighing step and proceed directly to **step 3** using 1 mL of TRIZOL reagent.
4. For fibrous tissues, such as rat and mouse heart and skeletal muscle, the most difficult step in the isolation process is the complete disruption of all cells when one is preparing tissue homogenates. Owing to low cell density and the polynucleated nature of cardiomyocytes, the yield of total RNA is typically low. In addition, fibrous tissue is difficult to homogenize completely, which can result in degraded RNA and very low yield. For complete lysis of the cells, keeping the tissue completely frozen until homogenization and using a powerful homogenizer

is critical to isolating intact total RNA. Embryonic heart tissue can be easily homogenized using a hand-held glass homogenizer.

5. Samples should not be thawed before contact with TRIZOL reagent to avoid degradation of the RNA.
6. The RNA should not be stored in water for a longer period to avoid degradation. For long-term storage, keep the RNA in an alcohol solution at –80°C. For this purpose, add 2 vol of 100% ethanol and 1/10 vol of 3 *M* sodium acetate (NaOAc), pH 5.2, to the RNA in water. Before use, spin down the RNA in a microcentrifuge at 15,000*g* for at least 15 min at 4°C, decant the supernatant, and wash the pellet with 1 mL 80% ethanol. Centrifuge again for 5 min, remove the supernatant, and dry the pellet for 10 min at RT.
7. At this point, we frequently do a "prerun" of the chip to make sure that the chip does not have bubbles. Give the chip a quick vortex (10 s, 512*g*), and proceed with the assay (running the chip) as usual. Allow the assay to run until it *focuses* for 10 s. If this is successful, abort the run, and proceed with adding the samples.
8. Agilent has recently introduced the RNA Integrity Number (RIN) as part of a software extension. The RIN is a tool to grade RNA quality on a quantitative scale of 1 (worst) to 10 (best). Sample integrity is determined by the entire electrophoretic trace of the RNA sample, including the presence or absence of degradation products.
9. Total RNA versus PolyA + RNA: good-quality microarray data have been obtained using total RNA or PolyA+ RNA as starting material for sample labeling. The results obtained from both types of sample are similar but not identical. Therefore, only samples prepared using the same sample preparation protocol should be compared. It is advisable to isolate total RNA first and check the initial RNA quality before proceeding with PolyA+ RNA isolation. PolyA+ RNA that has been purified once using oligo-dT columns might still contain significant amounts (up to 50%) of ribosomal and other nonpolyadenylated RNAs. Two rounds of purification over oligo-dT columns usually yield up to 95% pure PolyA+ RNA but result in significant loss of material. Therefore, PolyA+ RNA isolation depends on the availability of a sufficient amount of starting material.
10. The Affymetrix GeneChip One-cycle target labeling protocol is recommended for starting amounts of 1 to 15 µg total RNA or 0.2 to 2 µg PolyA RNA. Usually, 5 to 8 µg of total RNA yield around 80 to 100 µg of biotin-labeled cRNA. For hybridization, 15 to 20 µg of biotin-labeled cRNA is needed for a standard GeneChip. One microgram of RNA will yield just sufficient amounts for preparation of one hybridization cocktail. Here, the quality of RNA has to be excellent to ensure that enough biotin-labeled cRNA is being made. If less than 1 µg of total RNA is available, an additional round of amplification is required to generate the 15 µg amount of biotin-labeled cRNA needed for hybridization. Although most amplification methods claim that they faithfully maintain relative RNA abundance, amplification does shorten the length of the resulting transcripts and introduces artifacts, especially if the detection probes on the microarray are not sufficiently 3' biased. Hence, it is recommended to keep amplification to a minimum.

11. The volume of the eluate should be carefully measured by pipeting. Recovery of 11 µL or less may be a signal that the cDNA has not been completely eluted from the column.
    a. Check whether any liquid is present above the rim of the membrane. If liquid is seen along the rim of the membrane, use a P10 or P2 pipetor with a filtered, pointed (nonbeveled) pipet tip to aspirate the liquid gently from the rim and then eject the liquid directly onto the membrane itself. Centrifuge at maximum speed for 1 min.
    b. If there is still volume missing, perform an additional centrifugation step at maximum speed for 3 to 5 min.
    c. If a. and b. fail to recover the missing volume, add the missing volume (up to 5 µL) of cDNA Elution Buffer directly onto the membrane and repeat **step 8**.
12. If you are starting with more than 8 µg of total RNA, use no more than 6 µL of the double-stranded cDNA from **Subheading 3.5.3.** per IVT reaction.
13. Do not extract biotin-labeled RNA with phenol-chloroform. The biotin will cause some of the RNA to partition into the organic phase. This will result in low yields.
14. The average size distribution after one round of amplification should center around 1500 nucleotides. Additional rounds of amplification result in smaller fragment sizes, which can lead to signal loss if the microarray is not sufficiently 3'-biased.
15. Fragmentation is a critical step in the target preparation. Overfragmentation can cause high background, low signal, and high 3':5' ratios. If this is the case, fragmenting a new aliquot of cRNA for hybridization may solve the problem.
16. For different starting amounts, adjust the dilution accordingly to arrive at the same proportion of the spike-in controls to the RNA sample. Do not pipet less than 2 µL volumes, as this might increase the pipeting error and lead to inconsistencies.
17. The Affymetrix user guide for *Eukaryotic Sample and Array Processing* states that, if necessary, at this point, the probe array can be stored at 4°C for up to 3 h. However, it has been found that probe array processing at different time points after hybridization introduces variability. Hence, we recommend processing the probe array immediately after hybridization, if possible.

References

1. Downey, T., Pevsner, J., Jeon, O.-H., Murillo, F. M., and Spitznagel, E. L. (2004) Estimating and removing batch effects from microarray data, in *7th International Meeting of the Microarray Gene Expression Data Society*, September 8–10, 2004, Toronto, ON, Canada.
2. Affymetrix Manuals and User's Guides, available for download from www.Affymetrix.com:
    a. Expression Analysis Technical Manual, which contains the following sections:
       Section 1: GeneChip Expression Analysis Overview
       Section 2: Eukaryotic Sample and Array Processing
       Section 3: Prokaryotic Sample and Array Processing
       Section 4: Fluidics Station Maintenance Procedures
       Section 5: Appendices

b. Affymetrix Fluidics Station 400 Manual
c. Affymetrix Fluidics Station 450 Manual
d. Hybridization Oven Manual
e. Data Analysis Fundamentals
f. GeneChip® Operating Software Manual
g. Microarray Suite User's Guide, Version 5.0

3. Irizarry, R. A., Bolstad, B. M., Collin, F., Cope, L. M., Hobbs, B., and Speed, T. P. (2003) Summaries of Affymetrix GeneChip probe level data. *Nucleic Acids Res.* **31,** e15.
4. Li, C. and Wong, W. H. (2001) Model-based analysis of oligonucleotide arrays: expression index computation and outlier detection. *Proc. Natl. Acad. Sci. USA* **98,** 31–36.
5. Choe, S. E., Boutros, M., Michelson, A. M., Church, G. M., and Halfon, M. S. (2005) Preferred analysis methods for Affymetrix GeneChips revealed by a wholly defined control dataset. *Genome Biol.* **6,** R16.
6. Knudtson, K. L., Griffin, C., Brooks, A., et al. (2002) Factors contributing to variability in DNA microarray results: the ABRF Microarray Research Group 2002 study. www.abrf.org.