GeneChip®
Eukaryotic Small Sample Preparation Technical Note

Introduction

The standard GeneChip® eukaryotic target labeling protocol has been used by Affymetrix® customers for expression analysis on GeneChip probe arrays[1] . However, since the standard procedure requires a recommended 5 µg of total RNA as starting material for each target preparation reaction, a different protocol is necessary in situations where much smaller amounts of RNA are available for expression profiling.

For researchers working with small samples such as small biopsies, laser microdissected tissues or flow-sorted cells, Affymetrix has developed a protocol that carries out two cycles of standard cDNA synthesis and *in vitro* transcription (IVT) for GeneChip target amplification.

To evaluate the performance of this protocol, targets prepared using this labeling procedure from mouse heart total RNA were hybridized to GeneChip® Mu11KsubA arrays. The cRNA yield, linearity of amplification,

[1] Incyte Genomics, Inc. has asserted that the use of this protocol within the United States infringes U.S. Patent Nos. 5,716,785 and 5,891,636. Affymetrix and Incyte Genomics, Inc. are currently litigating this and related issues. Use of this protocol outside the United States does not infringe these patents. To the extent that users of this protocol wish to seek a license of these patents for use in the United States, they may contact Incyte Genomics, Inc., 3160 Porter Drive, Palo Alto, California 94304.

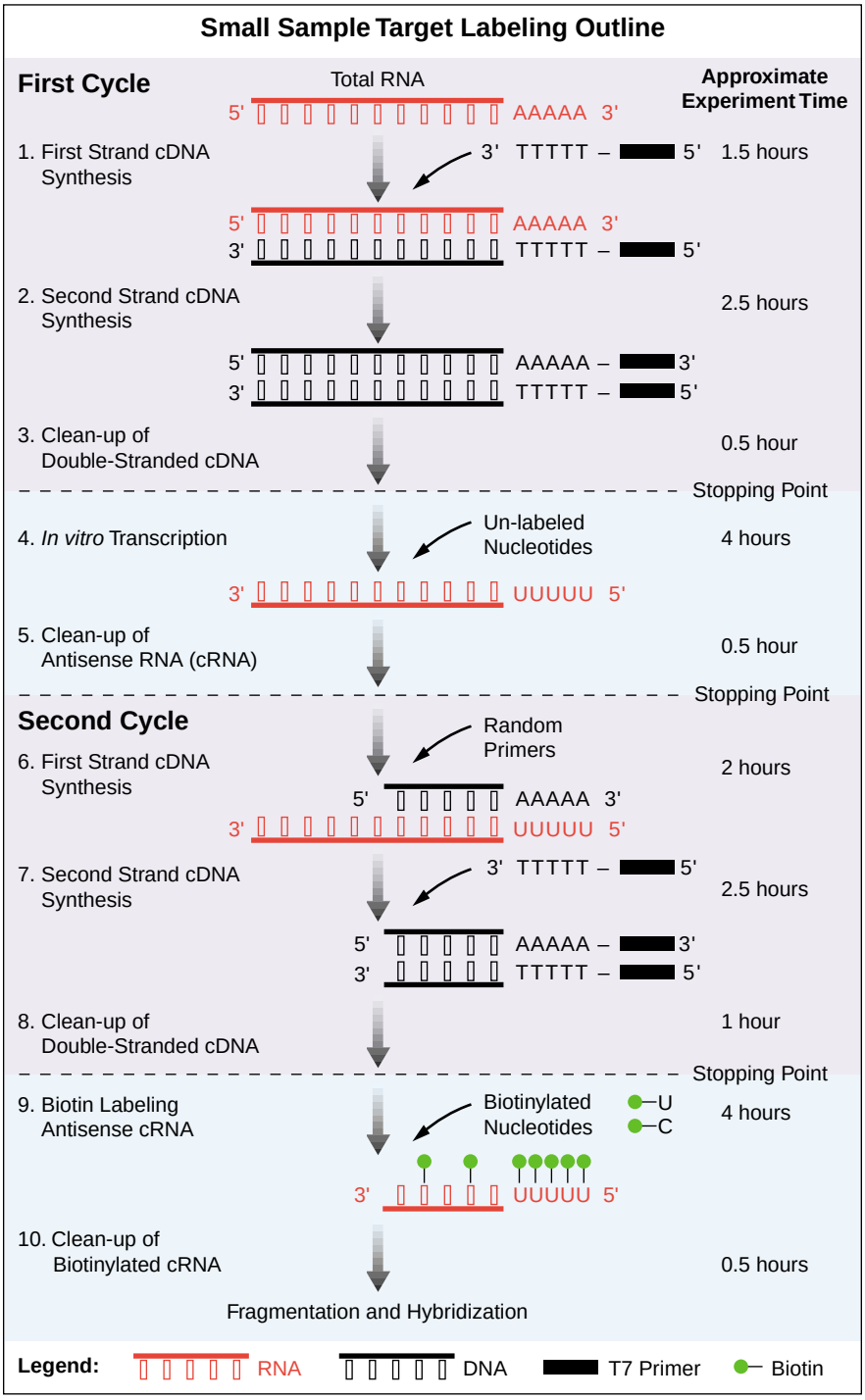

Figure 1. Schematic representation of small sample target labeling protocol.

assay reproducibility and representation of transcripts amplified were examined. Both the detailed experimental procedure and the data obtained on GeneChip arrays are described in this Technical Note.

It is important to note, however, that this research protocol has not been fully tested and validated on GeneChip probe arrays; therefore, we are not able to include recommendations under variable conditions and technical support from Affymetrix is limited. In the process of evaluating different assay strategies, we are interested in sharing with customers in this Technical Note some of our initial research observations. The procedures discussed herein should be considered recommendations only. If you choose to follow and test the protocol below, it should first be validated in your laboratory.

Assay Principles

The small sample target labeling protocol utilizes two cycles of cDNA synthesis combined with *in vitro* transcription for target amplification as outlined in Figure 1 (Eberwine *et al.*, 1992). The first cycle provides initial amplification of total RNA, resulting in unlabeled cRNA. In the second cycle of IVT synthesis, biotin-ribonucleotides are incorporated to produce labeled antisense cRNA target. The reagents and materials used in the assay are listed in Appendix 1 and the detailed protocol is described in Appendix 2.

Results

Several parameters were used to evaluate the robustness of the small sample target labeling technique for microarray expression analysis. The results were compared with standard GeneChip® expression analysis protocol including:

- **Yield of labeled cRNA**
- **Representation in amplifying various transcripts**
- **Linearity of amplification**
- **Sensitivity and reproducibility**

All results presented in this Technical Note were obtained on arrays that were stained with the Single Stain protocol and scanned on a GeneArray® scanner with the unadjusted PMT setting. Similar results were reproduced using the Antibody Amplification Staining protocol under the recalibrated PMT setting.

Yield of Labeled cRNA

The cRNA yield after two cycles of amplification from 100 ng, 50 ng, 20 ng, 10 ng and 1 ng of mouse heart total RNA was compared with the cRNA yield from 10 µg of the same RNA sample, using the standard protocol. Replicate experiments were carried out; the quantities of cRNA obtained are shown in Figure 2. The cRNA yield from all starting sample sizes was repeatedly above 10 µg, sufficient for hybridization to up to five GeneChip® probe arrays.

Representation in Amplifying Various Transcripts

To evaluate whether the small sample target labeling procedure produced high-quality cRNA, 10 µg of labeled cRNA target was hybridized to GeneChip® Mu11KsubA arrays. The data were analyzed with Microarray Suite 4.0 software. Replicate targets were prepared using either the standard protocol from 10 µg of total RNA, or using the small sample protocol from 100 ng, 50 ng, 20 ng, 10 ng and 1 ng of the same sample RNA. The percentage of probe sets in each experiment called Present is shown in Figure 3 (blue circles) and the average percentage Present call of replicates is shown as the red bar. The 100-ng and 50-ng samples resulted in

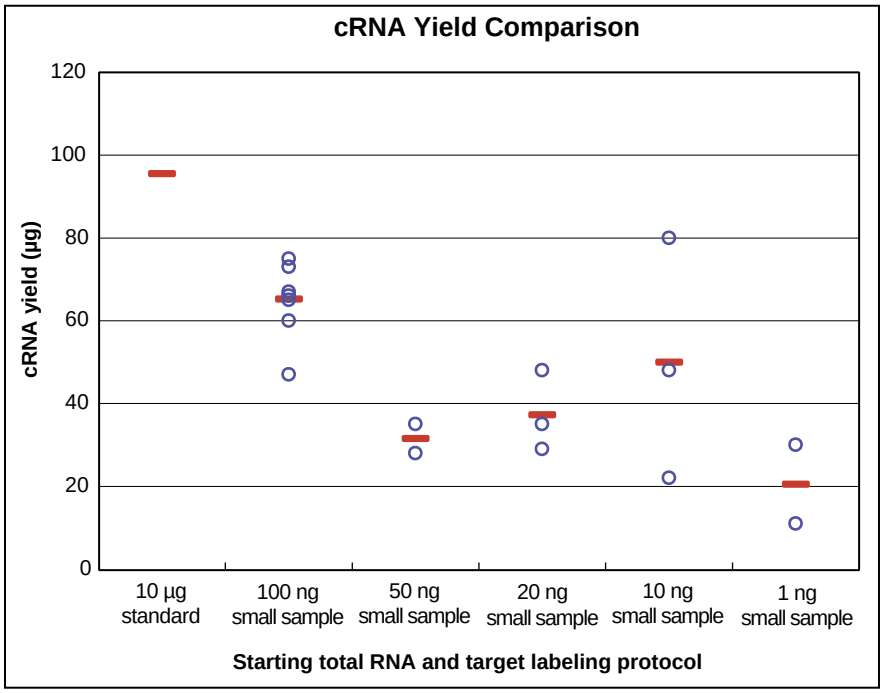

Figure 2. Comparison of cRNA yields. The cRNA targets were prepared using either the standard GeneChip® assay (for the 10-µg total RNA sample from mouse heart) or the small sample target labeling protocol (for 100 ng, 50 ng, 20 ng, 10 ng and 1 ng of the same sample). The cRNA obtained was quantitated with absorbance at 260 nm. Yields from individual replicates are plotted (blue circles) to demonstrate the repeatability of amplification reactions. Average yield of replicate samples is shown with the red bar.

approximately 44% Present calls, comparable to that routinely obtained from 10 µg of starting total RNA labeled with the standard protocol. A gradual decrease in percentage Present calls started with the 20-ng samples and only 20% of the probe sets were called Present with 1 ng of total RNA.

To more closely examine the reproducibility of the assay and concordance of detectable transcripts amplified by the two different protocols, we analyzed the extent of overlap of probe sets categorized as Present or Absent between different experiments. We first determined the expected reproducibility/concordance between two replicate samples using the same protocol. We did so in order to establish the baseline prior to comparing samples processed with different protocols. For this purpose, two replicates each of the 10-µg samples and 100-ng samples were used, and the discordant calls between the two replicates are plotted. As demonstrated in Figure 4, for both 10-µg and 100-ng replicate samples, approximately 5% of probe sets were called Present in replicate 1, but Absent in replicate 2. A similar percentage of discordant calls were also observed for the reverse (Absent in replicate 1 and Present in replicate 2). This indicates that the small sample target labeling assay (with 100-ng samples) performed as well as the standard expression assay (with 10-µg samples), with respect to absolute call concordance between replicates. Both protocols provided over 90% reproducibility.

One main concern with any protocol based on an amplification strategy is that the procedure may selectively amplify only certain transcripts, thus introducing a bias to the results.

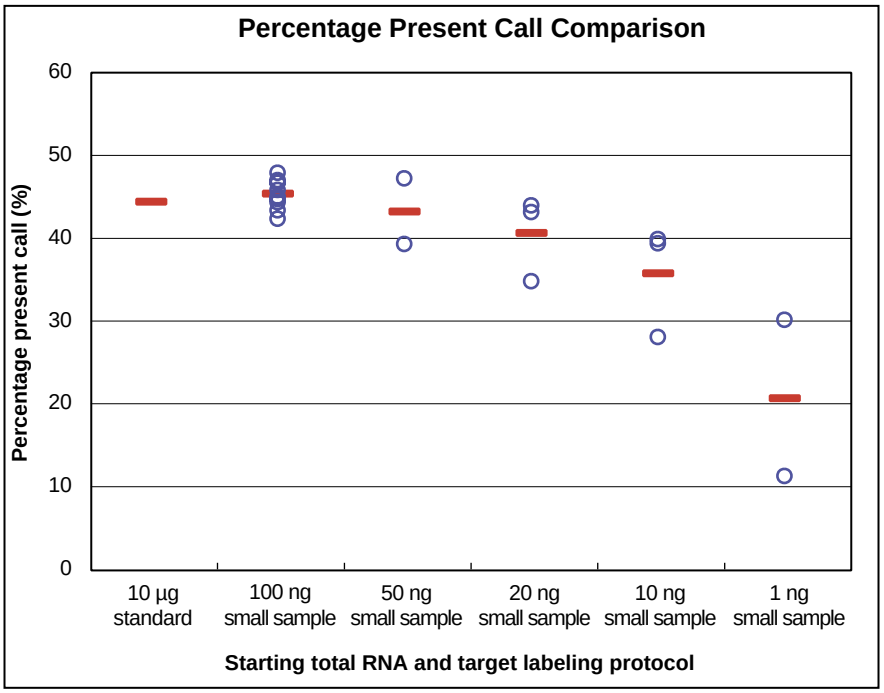

Figure 3. Comparison of percentage Present calls. Target generated from either the standard GeneChip® assay (for 10 µg of total RNA sample from mouse heart) or the small sample target labeling protocol (for 100 ng, 50 ng, 20 ng, 10 ng and 1 ng of the same sample) was hybridized to GeneChip® Mu11KsubA arrays. The percentage Present calls from individual arrays are plotted using blue circles to demonstrate the repeatability of the experiment, while the average percentage Present calls of replicate samples are shown in red bars.

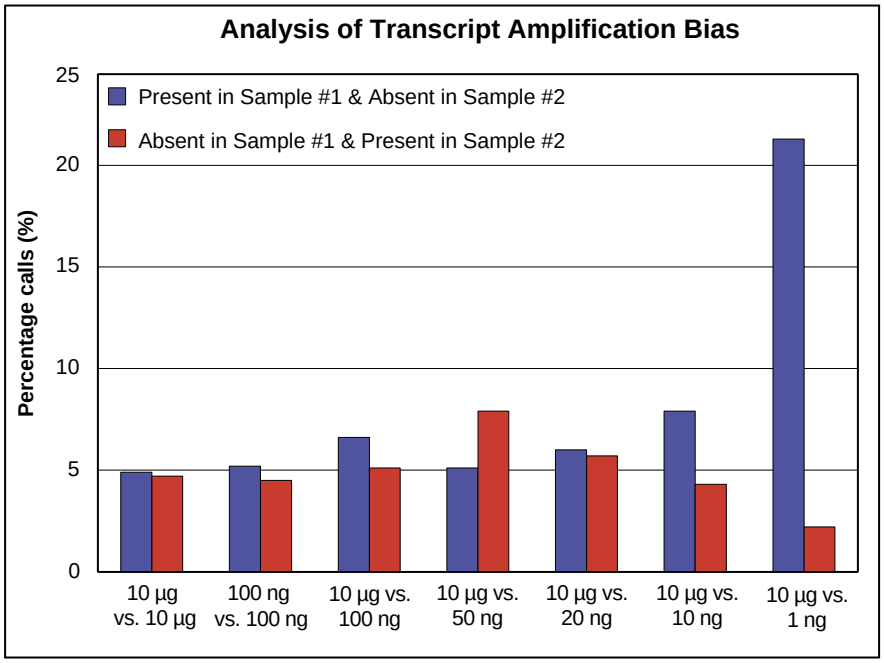

Figure 4. Analysis of transcript amplification bias. Target generated from either the standard GeneChip® assay (for 10 µg of total RNA sample from mouse heart) or the small sample target labeling protocol (for 100 ng, 50 ng, 20 ng, 10 ng and 1 ng of the same sample) was hybridized to GeneChip® Mu11KsubA arrays. Discordant calls represented as percentage of probe sets called Present in sample 1 and Absent in sample 2 (blue) or percentage of probe sets called Absent in sample 1 and Present in sample 2 (red) are plotted. Data from one comparison experiment are shown, but representative of other replicated results.

One way to evaluate possible bias of this protocol is to identify whether probe sets called Present in a standard assay from 10 µg of total RNA are also called Present with much reduced starting materials after two cycles of amplification. If the small sample target labeling protocol amplifies all transcripts as proportionally as the standard assay, we would anticipate a high degree of overlap of probe sets called Present from targets prepared by the two methods. Conversely, if the small sample target labeling assay amplifies only a subset of transcripts effectively, a much higher percentage of probe sets called Present by the standard assay may be called Absent by the small sample labeling assay since some transcripts may have dropped off. As shown in Figure 4, this particular category of discordant calls (blue bars) remains relatively constant at around 5%, similar to replicates performed using the same protocol, with as little as 10 ng of starting material. However, with very low amounts of starting material at 1 ng of total RNA, there is a decrease in Present calls indicating the procedure did not represent all transcripts in the sample. Therefore, this protocol may require further optimization for expression analysis with 1 ng or less of total RNA.

If the small sample target labeling protocol amplifies transcripts non-specifically, we would anticipate an increase in the percentage of probe sets called Absent by the standard assay, but Present by the small sample method. As shown in Figure 4, the probe sets called Present with the small sample protocol, but Absent with the standard protocol, also remained constant at around 5% with as low as 10 ng of starting total RNA (red bars).

In addition to comparing absolute calls obtained from the two labeling technologies, the degree of possible bias of the small sample target labeling protocol was also assessed by data concordance relative to the standard protocol. As discussed below, although direct comparison of Average Difference values between the two methods is not recommended, the concordance value does provide an indication whether there is severe amplification bias in an assay. The r^2 value for two replicate 100-ng samples labeled with the small sample protocol was above 0.96, similar to two samples processed with the standard procedure. An r^2 of ~0.88 was obtained when comparing intensities for samples as low as 50 ng using the small sample labeling protocol to 10-µg samples processed using the standard protocol. This indicated good concordance between the two protocols.

Linearity of Amplification

For quantitative expression analysis, it is crucial to optimize the amplification procedure so signals detected in a microarray experiment linearly reflect the concentrations of transcripts in the complex sample. To demonstrate that the small sample target labeling protocol maintains linearity of the assay, we spiked pre-determined amounts of exogenous poly(A)$^+$ transcripts into the sample at various concentrations. The samples were then labeled with either the standard GeneChip® assay or the small sample target labeling protocol. Intensity signals observed after hybridization to GeneChip® Mu11KsubA arrays are compared in Figure 5. With 10 µg or 100 ng of starting total RNA, the signals of spike-in controls were not distinguishable and closely related to concentrations at which the transcripts were spiked in.

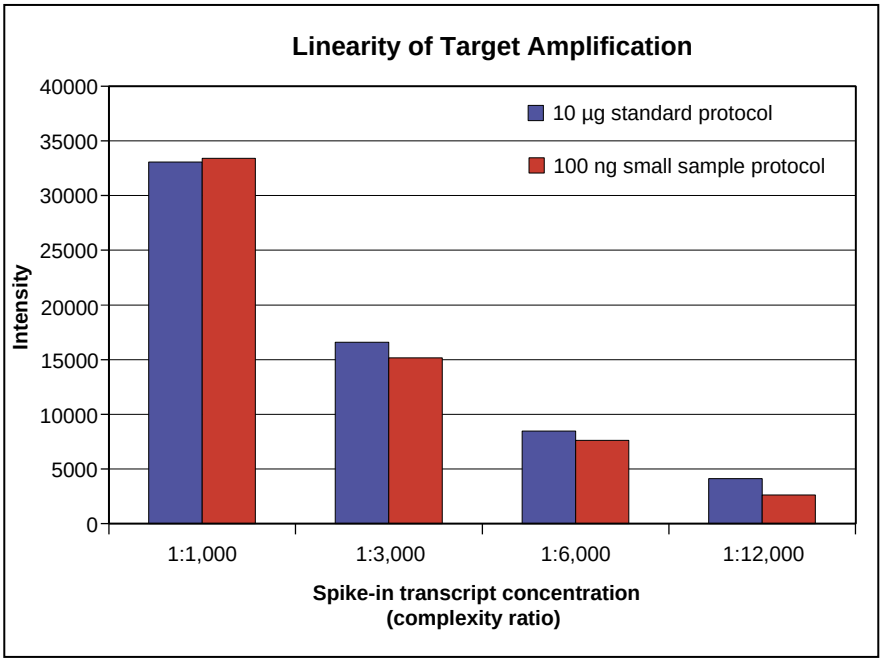

Figure 5. Assessment of amplification linearity. *Thr*, a *B. subtilis* gene, was used as the spike-in control. Poly-adenylated sense *thr* transcript was synthesized by *in vitro* transcription. After careful quantitation, the control poly(A)$^+$ transcript was spiked into the sample mouse total RNA at concentrations indicated in Figure 5. After scaling, intensities detected by the 3' probe set of *thr* are plotted. The complexity ratio was estimated based on the assumption that average length of a transcript in the complex sample is approximately 2,000 bases. The 1:12,000 complexity ratio in the mouse sample is roughly equivalent to 25 copies per cell.

Discussions

This small sample target labeling protocol was developed to amplify and label targets from limited total RNA samples for GeneChip® probe array expression analysis. Suitability of the protocol for this purpose has been examined using a range of starting concentrations of total RNA . Parameters characterized include labeled cRNA yield, percentage Present Calls, reproducibility, representation and linearity of amplification, as well as concordance with standard GeneChip assay protocol. Consistent results were obtained with as low as 50-100 ng of total RNA using this procedure.

The sensitivity of the assay was estimated by spiking in various concentrations of exogenous poly(A)$^+$ transcripts (data not shown). We only examined probe sets corresponding to the 3' end of genes due to 3' bias of the assay (see below). Based on limited experiments, we found the transcripts spiked in the total RNA at approximately 1:100,000 complexity ratio were called Present when amplifying 100 ng of total RNA, using the small sample target labeling protocol. This result was comparable to that obtained using 10 µg of starting material, labeled with the standard protocol. However, since detection of a given transcript is dependent on the probe set and the sequence from which the probes are selected, these inherent differences among probe sets may lead to variability in sensitivity for different transcripts.

Since the second cycle of amplification generates shorter products towards the 3' end of mRNA, as illustrated in Figure 1, we anticipate a stronger 3' bias of this protocol compared with the standard GeneChip expression assay. This is consistent with our observations of shorter labeled targets, when examining the cRNA on agarose gel, as well as higher 3' to 5' ratio with control genes including GAPDH. For example, we routinely obtained 3' to 5' ratio of between 2 and 4 for the small sample labeling procedure, compared to a ratio of between 1 and 2 for the standard assay.

In addition, since distinct targets are generated from a given total RNA using either the standard assay or the small sample target labeling assay, we find that although each assay itself is highly reproducible, it is not straightforward to compare signal intensities obtained from the two different methods. We do not recommend comparing results directly from one assay to another.

Reference

Eberwine, J., Yeh, H., Miyashiro, K., Cao, Y., Nair, S., Finnell, R., Zettel, M. and Coleman, P. (1992) *Proc. Natl. Acad. Sci. USA* **89:** 3010-3014.

Appendix 1: Reagents and Materials

REAGENT NAME	VENDOR	PART NUMBER
DEPC-treated water, 4 L	Ambion	9920
Total RNA, Mouse heart	Ambion	7816
T7-$(dT)_{24}$ Primer, HPLC-purified 5'- GGCCAGTGAATTGTAATACGACTCACTATAGGGAGGCGG - $(dT)_{24}$ - 3'	Genset	
SuperScript II, 200 U/µL, 40,000 U 5X First strand buffer and 0.1 M DTT included	Invitrogen	18064-071
dNTP mix, 10 mM, 100 µL	Invitrogen	18427-013
RNase inhitibor, 40 U/µL, 10,000 U	Ambion	2684
5X Second strand buffer, 500 µL	Invitrogen	10812-014
DNA Ligase, E.coli, 10 U/µL, 100 U	Invitrogen	18052-019
DNA Polymerase, *E.coli*, 10 U/µL, 1000 U	Invitrogen	18010-025
RNase H, 2 U/µL, 120 U	Invitrogen	18021-071
DNA Polymerase, T4, 5 U/µL, 250 U	Invitrogen	18005-025
Glycogen, 5 mg/µL, 5,000 µL	Ambion	9510
Ammonium Acetate (NH_4OAc), 5 M	Ambion	9070G
MEGAscript T7 Kit	Ambion	1334
RNeasy Mini Kit	Qiagen	74104
ß-Mercaptoethanol	Sigma	M-3148
Random primers	Invitrogen	48190-011
Enzo BioArray HighYield RNA Transcript Labeling Kit, 10 reactions	Affymetrix	900182
MISCELLANEOUS REAGENTS:		
Ethanol, absolute, kept cold		
Ethanol, 70%, kept cold		

Appendix 2: Methods

To evaluate the efficiency of the protocol, we carried out experiments with 100 ng, 50 ng, 20 ng, 10 ng and 1 ng of total RNA from mouse heart using the procedure described below. Targets were then hybridized to GeneChip® Mu11KsubA arrays. For comparison and concordance studies, 10 µg of the same mouse heart sample was also labeled using the standard assay protocol as detailed in the GeneChip® Expression Analysis Technical Manual.

Note: For experiments described in this Technical Note, we used aliquots of a single source of total RNA purified in bulk. Since GeneChip® expression analysis requires high-quality and high-purity RNA as starting materials, some general guidelines for RNA purification and quality assessment are provided in the GeneChip® Expression Analysis Technical Manual. However, we have not validated this protocol with RNA samples isolated directly from limited tissues or cells. Users may anticipate some additional optimization due to variability in quality of purified RNA.

First Cycle of Amplification

Step 1. First cycle, first strand cDNA synthesis

1. The sample RNA and T7-$(dT)_{24}$ primer were mixed in a microfuge tube:

	Volume/reaction	Final concentration
DEPC-treated water	9 µL	
Total RNA, variable concentration	1 µL	variable, ≤ 100 ng
T7-$(dT)_{24}$ primer, 100 pmol/µL	1 µL	100 pmol/reaction
Total volume	**11 µL**	

2. The RNA/primer mix was denatured by incubation at 70°C for 10 minutes.
3. Reaction mix was cooled on ice for 2 minutes.
4. The following reagents were added to the above RNA/primer mixture for the first strand cDNA synthesis:

	Volume/reaction	Final concentration
RNA/primer mixture	11 µL	
5X first strand buffer	4 µL	1X
DTT, 0.1 M	2 µL	10 mM
dNTP mix, 10 mM	1 µL	500 µM
RNase inhibitor, 40 U/µL	1 µL	40 U/reaction
Total volume	**19 µL**	

5. The reaction was incubated at 42°C for 2 minutes.
6. One microliter of SuperScript II (200 U/µL) was added to the above reaction to make a final volume of 20 µL.
7. The reagents were mixed gently and spun in a microcentrifuge briefly before incubating at 42°C for 1 hour.

Step 2. First cycle, second strand cDNA synthesis

1. The first strand reaction was spun down and cooled on ice.
2. The following reagents were added to the reaction tube for the second strand cDNA synthesis:

	Volume/reaction	Final concentration
1st strand synthesis reaction	20 µL	
DEPC-treated water	91 µL	
5X second strand buffer	30 µL	1X
dNTP mix, 10 mM	3 µL	200 µM
DNA ligase, *E. coli*, 10 U/µL	1 µL	10 U/reaction
DNA polymerase, *E. coli*, 10 U/µL	4 µL	40 U/reaction
RNase H, 2 U/µL	1 µL	2 U/reaction
Total volume	**150 µL**	

3. The reaction was mixed gently and spun in a microcentrifuge briefly followed by incubation at 16°C for 2 hours.
4. To fill in the ends of the double-stranded cDNA, 2 µL of T4 DNA polymerase (5 U/µL) was added to the above reaction and incubated at 16°C for an additional 15 minutes.

Step 3. First cycle, double-stranded cDNA clean-up by ethanol precipitation

1. To the reaction tube, 1 µL of 5 mg/mL glycogen, 0.6 volume of 5 M NH_4OAc (92 µL) and 2.5 volumes of cold absolute ethanol (612 µL) were added.
2. The content was mixed thoroughly and centrifuged immediately at 14,000 rpm for 20 minutes at 4°C.
3. The pellet was washed by adding 1 mL of 70% cold ethanol and centrifuged at 14,000 rpm for 5 minutes at 4°C.
4. After removing the ethanol carefully, a speed vacuum was used for approximately 5-10 minutes to dry the pellet. Avoid overdrying of the DNA. The pellet was stored at 4°C or -20°C overnight before proceeding to the IVT reaction.

Step 4. First cycle, IVT for cRNA amplification using Ambion MEGAscript T7 Kit

1. The following reagents were added to the above dried double-stranded cDNA pellet at room temperature as described in the Ambion MEGAscript protocol:

	Volume/reaction	Final concentration
DEPC-treated water	8 µL	
ATP, 75 mM	2 µL	7.5 mM
GTP, 75 mM	2 µL	7.5 mM
CTP, 75 mM	2 µL	7.5 mM
UTP, 75 mM	2 µL	7.5 mM
10X reaction buffer	2 µL	1X
Ambion Enzyme mix	2 µL	
Total volume	**20 µL**	

2. The reaction was mixed gently and spun in the microcentrifuge tube briefly before incubating at 37°C for 4 hours.

Step 5. First cycle, cRNA clean-up with RNeasy columns

1. Eighty microliter of RNase-free water was added to the above cRNA product.
2. The RNeasy Mini column was used for cRNA purification following the Protocol for RNA Cleanup handbook from Qiagen that accompanies the RNeasy Mini Kit.
3. In the last step of cRNA purification, the product was eluted with 50 µL of RNase-free water once.
4. The cRNA yield was determined by measuring the absorbance at 260 nm. Due to the limited amount of cRNA, the measurement may not be accurate. However, the rough estimation may help to determine the amount of cRNA to be used in the second cycle. The cRNA could be stored at -20°C at this point before proceeding to the next step.

Second Cycle of Amplification

Step 6. Second cycle, first strand cDNA synthesis

1. The following table was used as a general guideline to determine the amount of cRNA to be used in the second cycle:

Total Starting Material	cRNA To Be Used for second Amplification
50 - 100 ng	100 - 250 ng
< 20 ng	all 50 µL

Note:

For the instances when the yield was less than that indicated in the table, all of the elution (50 µL) was used in the second cycle of cDNA synthesis. For the reactions where the volume of cRNA needed exceeded 10 µL, a speed vacuum was used to reduce the volume to 10 µL before proceeding to the next step.

2. The cRNA and random primers were mixed in a microcentrifuge tube:

	Volume/reaction	Final concentration
cRNA, variable	up to 10 µL	variable
Random primers, 1 µg / µL	1 µL	1 µg/reaction
DEPC-treated water	add to final volume of 11 µL	
Total volume	**11 µL**	

3. The RNA was denatured by incubation at 70°C for 10 minutes.

4. The cRNA/primer mix was cooled on ice for 2 minutes.

5. The following reagents were added to the above cRNA/primer mixture for the first strand cDNA synthesis:

	Volume/reaction	**Final concentration**
cRNA/primer mixture	11 µL	
5X first strand buffer	4 µL	1X
DTT, 0.1 M	2 µL	10 mM
dNTP mix, 10 mM	1 µL	500 µM
RNase inhibitor, 40 U/µL	1 µL	40 U/reaction
Total volume	**19 µL**	

6. The reaction was incubated at 42°C for 2 minutes.

7. One microliter of SuperScript II (200 U/µL) was added to the above reaction to make a final volume of 20 µL.

8. The reagents were mixed gently and spun down in the microcentrifuge tube briefly before incubating at 42°C for 1 hour.

9. One microliter of RNase H (2 U/µL) was added and incubated for 20 minutes at 37°C.

10. The reaction was heated to 95°C for 5 minutes to denature the RNase H and separate the DNA/RNA hybrids.

11. The reaction was then chilled on ice.

Step 7. Second cycle, second strand cDNA synthesis with T7-$(dT)_{24}$ primer

1. The first strand reaction was spun down in a microcentrifuge, then 1.0 µL of 100 pmol/µL T7-$(dT)_{24}$ primer was added.

2. The mix was incubated at 70°C for 10 minutes and then cooled on ice.

3. The following reagents were added to the reaction tube for the second strand cDNA synthesis:

	Volume/reaction	**Final concentration**
First strand mix/ T7-$(dT)_{24}$ primer	22 µL	
DEPC-treated water	91 µL	
5X second strand buffer	30 µL	1X
dNTP mix, 10 mM	3 µL	200 µM
DNA polymerase, *E. coli*, 10 U/mL	4 µL	40 U/reaction
Total volume	**150 µL**	

4. The reaction was mixed gently and spun in the microcentrifuge tube briefly before incubating at 16°C for 2 hours.

5. To fill in the ends of the double-stranded cDNA, 2 µL of T4 DNA polymerase (5 U/µL) was added to the above reaction and the reaction was incubated at 16°C for an additional 15 minutes.

Step 8. Second cycle, double-stranded cDNA clean-up by ethanol precipitation

1. To the reaction tube, 1 µL of 5 mg/mL glycogen, 0.6 volume of 5 M NH_4OAc (92 µL) and 2.5 volumes of cold absolute ethanol (612 µL) were added.

2. The content was mixed thoroughly and centrifuged immediately at 14,000 rpm for 20 minutes at 4°C.

3. The pellet was washed by adding 1 mL of 70% cold ethanol and centrifuged at 14,000 rpm for 5 minutes at 4°C.

4. Ethanol was removed carefully and put in a speed vacuum for approximately 5 - 10 minutes to dry the pellet. The pellet was stored at 4°C or -20°C overnight before proceeding to the IVT reaction.

Step 9. Second round, IVT for cRNA amplification and labeling with Enzo BioArray HighYield RNA Transcript Labeling Kit

1. The following reagents were added to the above dried double-stranded cDNA pellet at room temperature as described in the Enzo kit product insert:

	Volume/reaction	Final concentration
DEPC-treated water	22 µL	
10X HY reaction buffer	4 µL	1X
10X Biotin labeled ribonucleotides	4 µL	1X
10X DTT	4 µL	1X
10X RNase inhitibor mix	4 µL	1X
20X T7 RNA polymerase	2 µL	1X
Total volume	**40 µL**	

2. The reaction was mixed gently and spun in the microcentrifuge tube briefly before incubating at 37°C for 4 hours.

Step 10. Second round, labeled cRNA target clean-up with RNeasy columns

1. Sixty µL of RNase-free water was added to the above cRNA product.
2. The RNeasy Mini column was used for cRNA purification. Protocol for RNA Cleanup handbook from Qiagen that accompanies the RNeasy Mini Kit was followed.
3. In the last step of cRNA purification, the product was eluted with 50 µL of RNase-free water once.
4. Two µL of the labeled cRNA was removed and added to 98 µL of water to measure the absorbance at 260 nm for determination of the cRNA yield.
5. Ten micrograms of labeled cRNA was then fragmented and hybridized to GeneChip® probe arrays as described in the GeneChip Expression Analysis Technical Manual.

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

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

www.affymetrix.com

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

Part No. 701129 Rev 1

AFFYMETRIX®