Digital Multiplexed Gene Expression Analysis Using the NanoString nCounter System

Meghana M. Kulkarni[1]

[1]Harvard Medical School, Boston, Massachusetts

ABSTRACT

This unit presents the protocol for the NanoString nCounter Gene Expression Assay, a robust and highly reproducible method for detecting the expression of up to 800 genes in a single reaction with high sensitivity and linearity across a broad range of expression levels. The methodology serves to bridge the gap between genome-wide (microarrays) and targeted (real-time quantitative PCR) expression profiling. The nCounter assay is based on direct digital detection of mRNA molecules of interest using target-specific, color-coded probe pairs. It does not require the conversion of mRNA to cDNA by reverse transcription or the amplification of the resulting cDNA by PCR. Each target gene of interest is detected using a pair of reporter and capture probes carrying 35- to 50-base target-specific sequences. In addition, each reporter probe carries a unique color code at the 5′ end that enables the molecular barcoding of the genes of interest, while the capture probes all carry a biotin label at the 3′ end that provides a molecular handle for attachment of target genes to facilitate downstream digital detection. After solution-phase hybridization between target mRNA and reporter-capture probe pairs, excess probes are removed and the probe/target complexes are aligned and immobilized in the nCounter cartridge, which is then placed in a digital analyzer for image acquisition and data processing. Hundreds of thousands of color codes designating mRNA targets of interest are directly imaged on the surface of the cartridge. The expression level of a gene is measured by counting the number of times the color-coded barcode for that gene is detected, and the barcode counts are then tabulated. *Curr. Protoc. Mol. Biol.* 94:25B.10.1-25B.10.17. © 2011 by John Wiley & Sons, Inc.

Keywords: gene expression signature • multiplex analysis • signal transduction • high-throughput screening • molecular barcode

INTRODUCTION

Virtually any biological condition—whether a developmental state, a cellular response to extracellular stimulus, or a pathological state—is reflected by changes in gene expression. One measure of the expression level of a gene is the number of corresponding mRNA molecules present in the cell. While the expression level of a single gene cannot define biological states, the ability to measure gene expression on a genome-wide scale using technologies such as microarrays provides an opportunity to identify gene expression signatures relevant to a particular biological phenomenon or response. Gene expression signatures are commonly composed of tens to hundreds of genes (Amit et al., 2009). The collective expression pattern of genes within a particular signature has been used in a number of studies to predict disease type and progression (e.g., van de Vijver et al., 2002; Burczynski et al., 2006). More importantly, gene signature-based analysis has the potential to define the molecular mechanism(s) by which genes contribute to a particular biological process, and can thus not only inform the choice of therapeutic intervention but also predict patient response to therapy (Lamb et al., 2003, 2006; Glas et al., 2005).

The ability to identify specific gene expression signatures that reflect different biological states raises the possibility of performing high-throughput small-molecule or RNA-interference screens using a signature of interest as the read out. Such screening would allow interrogation of pharmacological and genetic perturbations to assess the extent to which they exhibit the same or distinct signatures (Lamb et al., 2006). Linking biological states with specific perturbations (genetic and/or pharmacological) using quantitative information from changes in the expression of corresponding signature genes can provide a global understanding of the organization of cellular signaling circuits. The success of this approach depends on the ability to screen a large number of genes and compounds in a timely and cost-effective manner. Therefore, quantitative procedures for high-throughput measurement of mRNA levels in a simple, flexible, and cost-effective manner are important.

This unit presents the protocol for the nCounter Gene Expression Assay (*http://NanoString.com/*), a robust and highly reproducible method for detecting the expression of up to 800 genes in a single reaction. The methodology serves to bridge the gap between genome-wide (microarrays) and targeted (real-time quantitative PCR) expression profiling by virtue of its ability to multiplex up to 800 gene transcripts within a single reaction, with high sensitivity and linearity across a broad range of expression levels. The nCounter assay is based on direct imaging of mRNA molecules of interest that are detected using target-specific, color-coded probe pairs (Geiss et al., 2008). It does not require the conversion of mRNA to cDNA via reverse transcription or the amplification of the resulting cDNA via PCR. The large dynamic range of the assay and the lack of any enzymatic reactions allows direct analysis of a wide variety of sample types, including purified total RNA, crude cell and tissue lysates, RNA extracted from formalin-fixed paraffin-embedded (FFPE) tissues, and blood collected without globin mitigation.

Each target mRNA of interest is detected using a pair of probes called the reporter and capture probes. Each probe within a probe pair contains target-specific sequences that together recognize a 100-base contiguous region within the targeted mRNA (Fig. 25B.10.1). In addition, the reporter probes carry a color code at the 5′ end that enables the molecular barcoding of the genes of interest, while the capture probes carry a biotin label at the 3′ end that provides a molecular handle for the attachment of target genes to facilitate downstream digital detection. The color code on each reporter probe has six positions, each of which can be one of four spectrally non-overlapping fluorescent colors. The different combinations of the four distinct colors at six contiguous positions allows for a large diversity of color-based barcodes, each designating a different gene transcript, that can be mixed together in a single reaction for hybridization and still be individually resolved and identified. It is theoretically possible to generate 4^6 (4,096) different color codes, each barcoding a distinct mRNA. However, the kinetics of solution-phase hybridization and the technology for digital detection of different color barcodes currently limit the measurement power to less than a thousand mRNA species.

The target mRNA is mixed in solution with a large excess of the reporter and capture probe pairs, so that each targeted transcript finds its corresponding probe pair. After hybridization, excess unbound probes are washed away and the tripartite complexes, comprising target mRNA bound to specific reporter-capture probe pairs, are isolated. The biotin label at the 3′ end of the capture probes is used to attach the complexes to streptavidin-coated slides. An electric field is applied to orient and extend the tripartite complexes on the surface of the slide to facilitate imaging and detection of the color-coded molecules. A microscope objective and a CCD camera are used to image the immobilized complexes. The number of molecules for a particular mRNA species is counted by decoding the unique pattern of the fluorescent colors encoded in each reporter probe. The protocol is performed from start to finish on the nCounter System, which

is designed to provide hybridization, post-hybridization processing, and digital data acquisition capabilities in one simple workflow. The integrated system is composed of two instruments: the fully automated nCounter Prep Station for post-hybridization processing and the Digital Analyzer for imaging, data collection, and data processing. The CodeSet containing the reporter and capture probes targeting the genes of interest as defined by the user and all the reagents and consumables needed to perform the analysis are provided ready to use.

STRATEGIC PLANNING

The nCounter assay can be used to detect many types of nucleic acids including mRNA, DNA, and microRNA. This unit describes the protocol for measuring mRNA levels (Fig. 25B.10.1). The entire procedure can be divided into four major sections: (1) hybridization of target mRNA to gene-specific probe pairs, (2) fully automated post-hybridization processing, (3) data acquisition (imaging) and processing of barcodes, and (4) data analysis. Here, a brief overview of the protocol and underlying principles is provided, followed by detailed discussions of the gene signature, CodeSet, controls, and sample preparation.

Overview of the nCounter Gene Expression Assay

A multiplexed CodeSet is assembled with two sequence-specific probes for each target gene of interest (Fig. 25B.10.1A). The capture probe contains, from 5′ to 3′, a 35- to 50-base sequence complementary to a particular target mRNA, a short sequence containing two repeats of a 15-base sequence common to all capture probes (orange segment in Fig.25B.10.1A), and a biotin affinity tag. The reporter probe contains, from 3′ to 5′, a second 35- to 50-base sequence (complementary to the same target mRNA, near or contiguous with the target-specific sequence in the capture probe partner), a short sequence containing four repeats of a 15-base sequence common to all reporter probes (black segment in Fig. 25B.10.1A), and a color-coded molecular barcode. The common sequences included in all capture and reporter probes facilitate the removal of unbound excess probes during post-hybridization processing. The barcode is composed of a linearized single-stranded M13 DNA molecule annealed to a series of six complementary RNA segments, each labeled with one of four spectrally non-overlapping fluorescent dyes. The linear arrangement of the differently colored RNA segments creates a unique color code for each target gene of interest. The expression level of each gene is determined by scoring the number of times its corresponding color code is detected. In addition to the target-specific probe pairs, the CodeSet contains unique pairs of capture and reporter probes for reference genes specified by the user, as well as a number of nCounter system controls, including eight negative controls and seven positive spike-in controls for which RNA targets are added to the reactions. The methodology offers the flexibility of multiplexing up to 800 reporter-capture probe pairs within a single reaction.

The CodeSet is mixed in large excess with the mRNA sample in a solution-phase hybridization reaction (Fig. 25B.10.1B). Hybridization results in the formation of tripartite complexes, each containing a target mRNA bound to its specific capture and reporter probes. Each total RNA sample is hybridized in triplicate, with each reaction containing ~150 ng total RNA, 200 pM of each capture probe, and 40 pM of each reporter probe. The hybridization reaction is carried out in a standard thermocycler for 16 hr.

Post-hybridization processing is carried out on a fully automated nCounter Prep Station liquid-handling robot (Fig. 25B.10.1C). First, unhybridized probes are sequentially removed using magnetic beads coupled to short oligonucleotides complementary to the common sequence tags on the capture and reporter probes (orange and black segments in Fig. 25B.10.1A). The hybridization reaction is first purified using beads with the

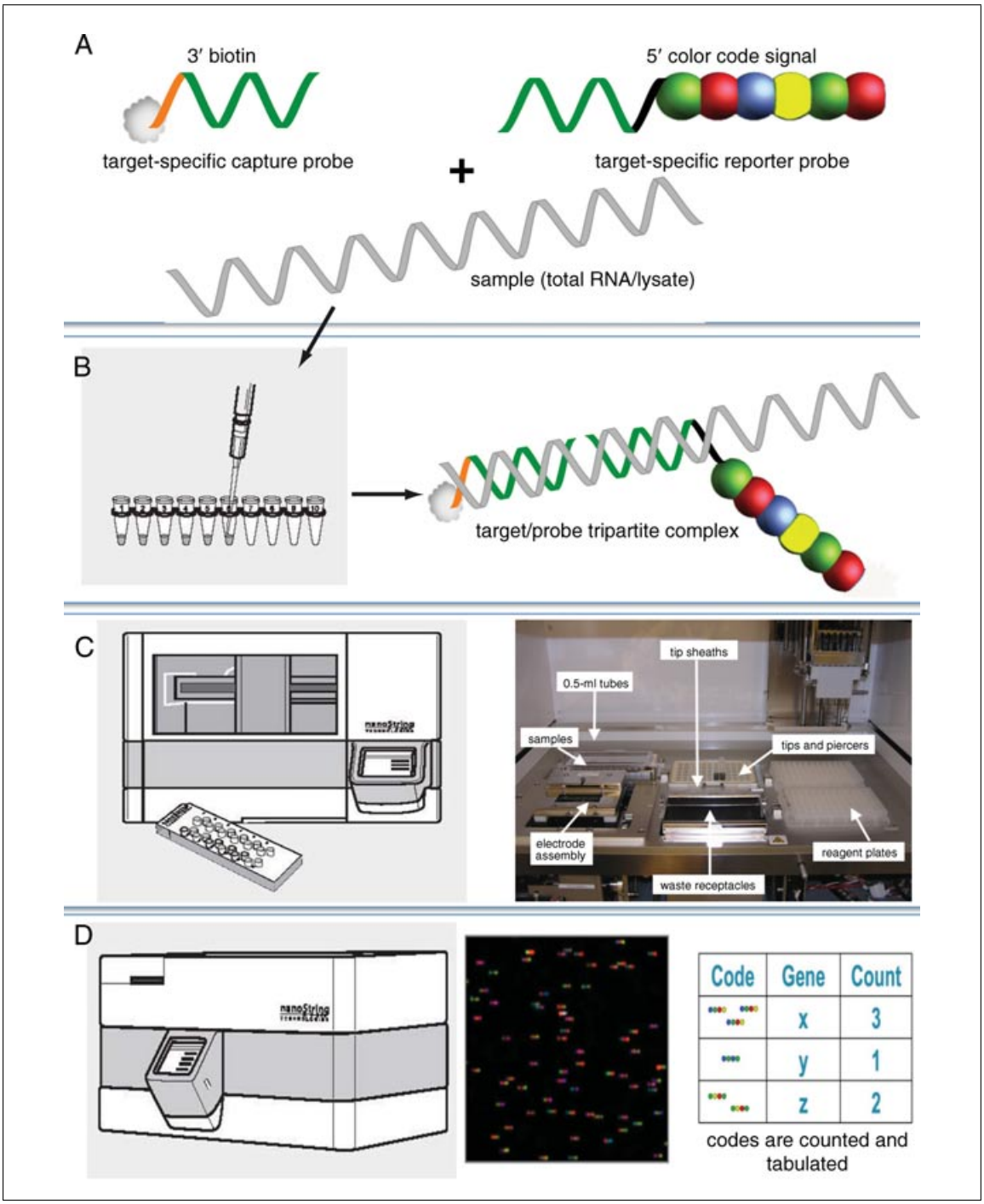

Figure 25B.10.1 Overview of nCounter Gene Expression Assay. **(A)** Schematic of target-specific capture and reporter probes, showing the 5′ reporter barcode and the 3′ biotin capture handle. The orange segment in the capture probe represents a 30-base sequence common to all capture probes, and the black segment in the reporter probe represents a 60-base sequence common to all reporter probes. Both are used to remove excess unbound probe. **(B)** During hybridization, buffer, the CodeSet (capture and reporter probe library) and sample (total RNA or cell lysate) are combined in strip-tubes and hybridized overnight at 65°C. Hybridization results in the formation of tripartite structures comprising target mRNA bound to specific capture and reporter probes. **(C)** Strip-tubes containing tripartite complexes are placed on the Prep Station (left) for automated post-hybridization processing. The deck layout is shown on the right. **(D)** The loaded sample cartridge is transferred to the Digital Analyzer (left) for imaging and data processing. A raw image of color-coded reporter probes bound to complementary target mRNA is shown (middle) along with a schematic of the tabulated counts for target genes (right). Modified with permission from NanoString Technologies.

common sequence of the capture probe, washed in a buffer containing low salt, and eluted off the beads at elevated temperature. This step removes excess unbound reporter probes, partial complexes containing only the reporter and target mRNA, and non-target cellular transcripts not bound to a capture probe. The eluate, containing capture probes and tripartite complexes, is then purified with a second round of beads with the common sequence of the reporter probe to remove excess unbound capture probes. Second, the purified complexes are loaded into the NanoString Sample Cartridge using hydrostatic pressure and bound to the streptavidin-coated surface via the 3′ biotin label of the capture probe. Each cartridge contains 30-μm-deep microfluidic channels that can be loaded with twelve different samples at a time. After attachment to the surface of the cartridge, the bound molecules are extended and aligned by applying an electric field (160 V/cm for 1 min) along each of the twelve flow cells. The extended complexes are further immobilized at the reporter probes by the addition of biotinylated oligonucleotides complementary to the 5′ end of each reporter probe. After immobilization, a SlowFade reagent is added to each flow cell prior to imaging.

The NanoString cartridges containing the immobilized complexes are sealed and imaged on the nCounter Digital Analyzer (Fig. 25B.10.1D), which takes images of each field of view (FOV) using four different excitation wavelengths (480, 545, 580, and 622 nm) corresponding to the four fluorescent dyes. At the highest resolution, images of each field of view are acquired at 60× magnification (600 fields of view) using a microscope objective and a CCD camera, yielding hundreds of thousands of counts of target mRNA molecules. The acquired digital images are processed on the Digital Analyzer. The expression level of a gene is measured by counting the number of times the specific barcode is detected, and the barcode counts are tabulated in a comma-separated value (CSV) format.

Selection of Gene Signature

In order to use changes in gene expression to link genes, drugs, and/or diseases to specific biological processes, it is imperative to establish useful expression signatures that robustly reflect the specific biological states of interest. An expression signature, which serves as a surrogate for a biological state of interest, can be defined as a set of genes whose pattern of expression is characteristic of that state. Thus, the first step is to define a set of genes whose pattern of expression can be used as readout for the process under study. This can be achieved by two general methods. Gene sets that accurately reflect a particular biological process can be determined empirically by genome-wide expression profiling using either microarrays or transcriptome sequencing by RNA-Seq (*UNIT 4.11*) under conditions where the biological process of interest is activated. For example, gene expression signatures that reflect activation of the insulin signaling pathway can be generated using RNA-Seq to measure genome-wide expression changes induced in cells stimulated with insulin at different time points. These gene expression data can then be used to identify a pattern of gene expression—a signature—that reflects the phenotype of insulin pathway activation. Amit et al. (2009) identified 118 signature genes that define the response of mouse primary dendritic cells to infection by pathogens using microarray expression profiling of dendritic cells exposed to different pathogen-derived components. For details of the statistical method developed to identify informative signature genes in this system, see Amit et al. (2009). For a general review of standard statistical methods applied to microarray data to derive gene expression signatures, see Nevins and Potti (2007). Once a robust gene expression signature representative of the biological process of interest is derived, one can interrogate pharmacological and genetic perturbations on a global scale via high-throughput small-molecule or RNA-interference screens to assess the extent to which these perturbations exhibit the same or distinct molecular phenotypes (Lamb et al., 2006). Such screening on a global scale using microarrays or RNA-Seq can be both technically challenging and prohibitively expensive.

Alternatively, gene sets that serve as surrogates for a biological state can be assembled from public databases. A number of public databases provide excellent resources for building useful gene expression signatures, including MSigDB (developed at the Massachusetts Institute of Technology and the Broad Institute; Subramanian et al., 2005) and the Oncomine cancer-profiling database (developed at the University of Michigan, Ann Arbor; Rhodes et al., 2004, 2005; Rhodes and Chinnaiyan, 2005). NanoString also offers pre-built gene expression panels and Virtual Gene sets (*http://www.nanostring.com/products/assays/*) that are composed of genes to study signaling pathways, disease states, and a number of other biological processes in human and mouse systems. While assembling a set of genes to represent the gene expression signature of a state of interest, it is important to select genes whose expression levels span a broad range of biological expression levels to minimize the contribution of noise inherent in hybridization-based assays across multiple experiments.

Design and Synthesis of CodeSet

Once a set of genes comprising the signature of interest has been established, the following four steps achieve the design and synthesis of gene-specific reporter and capture probes that will define the CodeSet. (1) The RefSeq (NCBI Reference Sequence database) accession numbers for the mRNA transcripts of the selected signature genes is submitted to NanoString using the CodeSet Design Form. (2) NanoString designs reporter and capture probe pairs for every gene in the signature gene list provided by the user. The NCBI Reference Sequence database is used as the source for the target sequences within each reporter and capture probe. To prevent off-target effects due to cross-hybridization of reporter-capture probe pairs to non-target molecules in the sample, the full non-redundant dataset containing the sequence information for a species from RefSeq is used to check the cross-hybridizing potential of the target-specific probe sequences. Probe pairs that target most, if not all, of the transcript variants for a particular gene are selected by default. Probes are screened for all standard hybridization parameters, such as hybridization efficiency, cross-hybridization potential, GC content, and predicted secondary structures. The best-scoring reporter-capture probe pair for each gene in the signature gene list is selected. A Design Report containing the scores for each hybridization parameter for the best reporter-capture probe pairs, including target sequence recognized by the probe pair, is sent to the user as an Excel file for final review. (3) The user reviews the Design Report. It is especially critical to cross-check that the sequence being recognized by the selected probe pairs corresponds to the sequence of the target genes selected by the user in the signature gene list. (4) NanoString completes the design and manufacture of the reporter and capture probe pairs approved by the user. The entire process, from submission of the list of signature genes to design and manufacture of the CodeSet containing reporter-capture probe pairs, is completed in 8 to 10 weeks.

Probe Replication and Control Probes

In any hybridization-based assay, it is critical to design several controls that can be used to determine whether the observed hybridization is specific to the target sequence. One approach is to use an alternate sequence (for example, the mismatch probes on an Affymetrix gene chip) or sense probes (rather than anti-sense) to distinguish between specific and non-specific (or off-target) hybridization. Another commonly used control is to design and use multiple probes for the same gene transcript. Given that most gene expression signatures are comprised of tens to several hundred genes, and that this approach makes it possible to multiplex up to 800 reporter-capture probe pairs in a single reaction, it is advisable to incorporate at least three independent probe pairs per target mRNA in the CodeSet design.

Gene expression assays employ one or two reference genes to normalize transcript levels between different samples, in order to account for differences in the amount of total RNA present in the samples. While these reference genes are often housekeeping genes, their expression levels frequently vary among tissues or cells, and may change under different conditions and treatments. Thus, the selection of appropriate reference gene(s) is critical for the success of gene expression studies. For this purpose, at least ten reference genes that do not show changes in expression levels (based on initial microarray analysis performed to select signature genes) should be chosen. It is preferable that the expression levels of the selected reference genes also span the dynamic range of the nCounter Gene Expression Assay (low, medium, and high expression levels). The assay has a linear dynamic range of 7×10^5 counts.

Positive spike-in controls (target sequences and corresponding reporter-capture probe pairs) are included in each CodeSet. Targets for the positive spike-in controls are 100-base oligonucleotides complementary to the spike-in reporter and capture probes, and are included in the reactions at final concentrations ranging from 0.1 to 100 fM. The specific 100-base regions are arbitrarily labeled POS_A to POS_G, and are generated from the following non-human sequences: NCBI RefSeq accession number AY058658.1 (A, E, F); accession number AY058560.1 (B-D); accession number DQ412624 (G).

In addition, the CodeSet includes eight negative control probe pairs (arbitrarily labeled NEG_A to NEG_H) for which there are no corresponding target sequences. These negative control sequences are part of a National Institute of Standards and Technologies (NIST) effort to generate a set of sequences that can be used to standardize across different gene expression assays (including microarrays and PCR-based platforms). These sequences, generated by the External RNA Control Consortium (ERCC) as part of the NIST, are "alien" in nature, i.e., they are synthetic or are from very obscure bacterial species and have been screened against the non-redundant sequence database for uniqueness. They should yield negative results in all common species used in research labs (Devonshire et al., 2010). The negative controls are used to calculate background hybridization signal and to determine present/absent calls for target genes.

Sample Preparation

A crucial component in a successful nCounter gene expression assay is the generation of high-quality total RNA. Total RNA from cells and tissues is generally extracted and purified using TRIzol Reagent (Invitrogen) according to the manufacturer's protocol. TRIzol-extracted total RNA should be further purified using RNeasy mini spin columns (Qiagen) according to the manufacturer's protocol. The concentration and quality of the total RNA preparation should determined by measuring absorbance at 260 and 280 nm using a standard UV spectrophotometer or the Nanodrop system (Thermo Scientific). Each hybridization reaction requires 100-150 ng of total RNA. The integrity of the total RNA preparation should be verified by denaturing PAGE or on a Bioanalyzer (model 2100, Agilent Technologies) before proceeding with the hybridization reaction. RNA should be aliquoted (to avoid freeze/thaw cycles) and stored for up to several years at $-80°C$.

Alternatively, hybridization can be set up using cell lysates. Cell lysates should be prepared using an RNeasy mini kit (Qiagen) according to manufacturer's recommendations, but using a concentration of 2,500 to 10,000 cells/μl of RLT buffer (Qiagen). The nCounter cell lysate hybridization protocol is optimized for ~10,000 mammalian cells/reaction, which is the equivalent of ~100 ng total RNA. Lysates should be used immediately after preparation to maintain the integrity of the cellular RNA.

For more detailed descriptions of the manufacturer's Total RNA Standard Protocol and Cell Lysate Protocol, see the nCounter Gene Expression Assay Manual at *http://www.nanostring.com/*.

BASIC PROTOCOL 1

HYBRIDIZATION OF TARGET mRNA TO GENE-SPECIFIC PROBE PAIRS

This protocol outlines the hybridization and post-hybridization procedures for measuring mRNA levels using the nCounter Gene Expression Assay. Hybridization is performed in tube-strips that are later loaded into a NanoString Sample Cartridge for post-hybridization processing. One cartridge can hold up to 12 different reactions (samples), so the user can set up 1 to 12 hybridization reactions at a time, keeping in mind that the cartridge must be discarded after use and cannot be re-used. Each sample should be hybridized in triplicate. The hybridization protocol should be followed exactly as described, as the order of addition of components is important.

When setting up the hybridization reactions, it is important not to mix reagents by vortexing or pipetting vigorously, as this may shear the reporter probes. Mixing should be done by gently flicking or inverting the tubes two to three times. If a microcentrifuge is used to collect the reaction mixture at the bottom of the tubes, do not spin at more than 1,000 rpm for more than 30 sec, and do not pulse spin, as this will cause the centrifuge to go to maximum speed, which may spin the CodeSet out of solution.

After hybridization, the nCounter Prep Station processes samples to prepare them for subsequent imaging and data processing. Post-hybridization processing includes removal of unbound reporter and capture probes, attachment of mRNA-reporter-capture complexes to streptavidin-coated Sample Cartridges, and extension and alignment of the complexes by application of an electric field. The Prep Station can process up to 12 samples per run in approximately 2 hr.

For a general description of the post-hybridization process, see Strategic Planning. For a detailed description of the Prep Station, including step-by-step images, maintenance, and troubleshooting information, refer to the nCounter Prep Station User Manual provided at *http://www.nanostring.com/*.

Materials

Sample (see Strategic Planning): total RNA (150 ng per hybridization reaction; −80°C) or cell lysate (from 2,500-10,000 cells; −80°C)
nCounter GX CodeSet (probe pairs for user-defined target genes; NanoString Technologies; −80°C)
Hybridization buffer (room temperature), provided in nCounter Prep Pack
RNase-free, DNase-free water

nCounter Prep Pack (NanoString Technologies), including: racked tips and foil piercers, 12-tube strips and caps, tip sheaths (for storing tips, to prevent cross-contamination and unnecessary tip consumption), and cartridge well seals (store at room temperature)
Tube-Strip PicoFuge (Stratagene)
Thermocycler or hybridization oven
nCounter Prep Station (NanoString Technologies, no. NCT-PREP-120)
nCounter Sample Cartridges (NanoString Technologies; store at −20°C)
nCounter Prep Plates (foil-sealed 96-well plates; NanoString Technologies; store at 4°C)
Benchtop centrifuge with swinging bucket rotor for plates (e.g., Eppendorf)

NOTE: Sample Cartridges, Prep Plates, and Prep Packs can be purchased together in a comprehensive nCounter Master Kit.

Perform hybridization

1. Remove cell lysate or total RNA samples and aliquots of reporter and capture probe reagents from the –80°C freezer and thaw on ice for 30 min. Invert several times to mix well and briefly spin down contents of the tubes.

2. Prepare a master mix containing 130 µl reporter probe and 130 µl hybridization buffer. Mix by gently flicking the tube and briefly spin down contents.

3. Dilute total RNA or cell lysate samples with RNase-free/DNase-free water to give a total RNA concentration of 30 ng/µl.

   For lysates, 10,000 mammalian cells is the equivalent of ~100 ng total RNA.

4. Label a 12-tube strip so that the sample in each tube is clearly indicated. Use triplicate tubes for each sample. Cut the tube-strips in half to fit in a picofuge.

5. Add 20 µl master mix to each of the 12 labeled strip-tubes.

   It is advisable to use a fresh tip for each pipetting step to accurately pipet the correct volume.

6. Add 5 µl diluted total RNA or cell lysate sample (containing 100-150 ng total RNA) to appropriately labeled triplicate strip-tubes containing master mix. Mix by gently flicking the tubes and briefly spin down contents using the tube-strip picofuge.

7. Pre-heat the thermocycler to 65°C. Program the thermocycler using a 30-µl volume, "calculated temperature" mode, heated lid, and "forever" time setting. Do not set the thermocycler to ramp down to 4°C at the end of the run.

   Alternatively, a 65°C hybridization oven may be used with a large beaker full of water placed in the oven to ensure a humid environment. It is ideal to maintain a constant temperature of 65°C during hybridization to ensure optimal reaction conditions. Avoid using a water bath or heat block, since evaporation over time could cause fluctuations in temperature.

8. Gently flick the tubes containing capture probe and briefly spin down contents. Add 5 µl capture probe set to each tube and immediately place tubes at 65°C. Cap tubes and mix the reagents by gently flicking the tube-strips several times. Briefly spin down contents and immediately place the tube-strip in the 65°C thermocycler.

   Minimizing the time between adding the capture probe and placing the reaction at 65°C will increase the sensitivity of the assay.

   The final hybridization reaction for each total RNA sample contains 5 µl sample RNA (150 ng), 10 µl reporter probe (final 40 pM), 5 µl capture probe (final 200 pM), and 10 µl hybridization buffer (5× SPPE, pH 7.5, with 0.1% Tween-20).

9. Incubate hybridization reactions at 65°C for at least 12 hr, optimally 16 hr. Maintain reactions at 65°C until ready for processing, but do not exceed 30 hr.

10. Once the hybridization reactions are removed from the thermocycler, proceed immediately to post-hybridization processing on the nCounter Prep Station. Do not store hybridization reactions.

Set up and run Prep Station

11. Turn on the nCounter Prep Station using the power switch located on the side of the machine.

    The nCounter Prep Station has a user-friendly touch screen with which the user can interface with the robot. Once powered, the Prep Station touch screen guides the user through the steps that must be followed to load the deck and start a run.

12. Bring a sealed Sample Cartridge and two Prep Plates to room temperature for at least 10 min before use.

    One Sample Cartridge and two Prep Plates are used to process 1 to 12 reactions in a single run. Prep Plates are 96-well plates containing all reagents for post-hybridization processing, including magnetic beads for removing unhybridized probe, wash buffer, and immobilization and imaging reagents.

To prevent accumulation of condensation on the cartridge, do not open the pouch seal until it has reached room temperature.

13. Collect contents of the Prep Plates in the bottom of the wells by centrifuging for 2 min at 670 × *g* in a standard benchtop centrifuge with a swinging bucket rotor.

14. Slide open the door of the Prep Station and empty the solid waste (containing tips) and liquid waste receptacles.

15. Briefly spin down the hybridized samples using a picofuge. Remove caps and place the hybridized sample tubes on the deck of the PrepStation, ensuring that tube 1 aligns with position 1.

    The numbering on the deck of the Prep Station corresponds to the numbers on the hybridized sample tubes. If the sample tubes are inverted, all sample-specific information will be assigned to incorrect data files.

16. On the Sample Selection touch screen, select all wells that are to be processed.

17. Remove the clear plastic lids from the Prep Plates and place the plates on the deck of the Prep Station as indicated on the touch screen, oriented so the labels face the user.

    Do not remove the foil seal or pierce the wells on the Prep Plates. The Prep Station pierces the wells during processing.

18. Remove the metal tip carrier from the Prep Station deck by lifting straight up. Place the tips and foil piercers into the carrier.

19. Replace the loaded tip carrier back onto the Prep Station deck so the foil piercers are closest to the user.

20. Place the tip sheaths on the deck and press firmly into place.

21. Place two strips of empty tubes (without caps) on the deck. Securely close the lid that flips down over the tubes.

22. Open the cartridge pouch and carefully place the cartridge under the electrode fixture in the correct orientation. Make sure that it is seated completely in the machined depression.

23. Carefully lower the electrode fixture in place over the cartridge so the 24 electrodes insert into the 24 wells. If any resistance is felt while lowering the fixture, stop and make sure that the electrodes are correctly aligned.

24. Close the door of the Prep Station by pulling down. Press Start on the touch screen to begin processing.

    The Prep Station first checks that all consumables and reagents have been placed in the correct orientation. To do this, the Prep Station confirms that the sensors for the sample cartridge, electrode fixture, and heater lid are all in the correct state. The pipet head then checks that tips, tip sheaths, strip tubes, and Prep Plates are all in place by touching them with a set of validation tips.

25. Upon completion of the run, remove the cartridge from the deck, taking care not to bend the electrodes. Seal the cartridge using cartridge well seals and label with a date and unique identifier.

    Target mRNA molecules with unique barcodes are now ready to be imaged, counted, and scored for relative abundance measurements.

DATA COLLECTION AND ANALYSIS

The nCounter Digital Analyzer acquires images of the immobilized fluorescently labeled target mRNA molecules in the sample cartridge using a CCD camera and a microscope objective lens. The analyzer takes images of each field of view (FOV) using four different excitation wavelengths (480, 545, 580, and 622 nm) corresponding to the four different fluorescent dyes in the barcode. At the highest resolution, images of each field of view are acquired at 60× magnification (600 fields of view) to yield hundreds of thousands of counts of target mRNA molecules. The acquired digital images are then processed internally on the analyzer. The expression level of a gene is measured by counting the number of times its specific barcode is detected, and the barcode counts are tabulated and reported in a CSV format. In this system, the number of reporter molecules counted determines the number of images acquired. This feature of the analyzer contributes in part to the large dynamic range and high sensitivity of the assay.

This protocol briefly describes the steps involved in image acquisition. For a more detailed step-by-step description of the process, refer to the nCounter Digital Analyzer User Manual provided at *http://www.nanostring.com/*.

Materials

nCounter Digital Analyzer (NanoString Technologies, no. NCT-DIGA-120)
nCounter Imaging Oil Applicator and optical oil, provided with Digital Analyzer
nCounter Memory Stick (NanoString Technologies), provided with CodeSet
Personal computer with Microsoft Excel

Collect data

1. Turn on the Digital Analyzer using the power switch located at the back of the machine.

   Similar to the Prep Station, the Digital Analyzer has a user-friendly touch screen interface that guides the user through the steps required to start image acquisition and data processing.

2. Apply optical oil to the undersurface of the sample cartridge.

   This step is important for achieving high-quality images. The optical oil is applied automatically in a stand-alone unit (the nCounter Imaging Oil Applicator) that accompanies the Digital Analyzer.

3. Open the door located at the top of the analyzer. Using a lens paper, wipe the objective clean prior to starting image acquisition. Place a drop of optical oil on the lens.

4. Place the cartridge in the correct orientation into one of the six stage positions. Be sure that the cartridge is seated flat in the slot and close the magnetic clips gently.

   The analyzer has six stage positions, each of which can be loaded with a cartridge.

5. Upload the reporter library file (RLF) specific to each CodeSet onto the analyzer. Save the RLF on the local analyzer drive.

   Each CodeSet is accompanied by a unique RLF, which is generated by NanoString and supplied on a memory stick. The RLF contains information (e.g., gene names, gene symbols, accession numbers) used during image processing to assign a target identity to each individual barcode.

   This step is performed only once for each user-defined CodeSet.

6. Create a cartridge definition file (CDF) and upload onto the local analyzer drive using the memory stick.

A CDF is created by the user for every cartridge analyzed. The CDF defines sample-specific data to associate with the data output. It also defines parameters used by the analyzer during image collection and processing, such as Lane ID (for each of the twelve flow cells in the cartridge), Sample ID (the name of the sample contained in each lane or flow cell), GeneRLF (the CodeSet RLF to associate with the data), and FOVcount (number of images to analyze per assay/flow cell). At the highest resolution, the FOVcount is set to 600 FOVs.

A CDF template is included on the memory stick and can be opened in Excel on the user's personal computer to create a CDF file for each cartridge.

7. Select Start Counting from the Main Menu on the touch screen.
8. Select the cartridge position corresponding to the stage where the sample cartridge has been loaded.
9. Select the cartridge-specific CDF that was saved to the local analyzer drive.
10. Shut the door of the instrument and select Start to begin image acquisition and data processing.

    The data are collected in a reporter code count (RCC) file created for each flow cell in the cartridge. Each RCC file contains the number of counts for each target mRNA molecule for one flow cell corresponding to one sample. A separate RCC file is created for each flow cell or sample in the cartridge. The twelve RCC data files corresponding to all samples in the cartridge are combined in an RCC.zip folder.

    At 600 FOV, the analyzer takes approximately 4-5 hr to acquire images for all twelve flow cells of a sample cartridge.

11. Select Download Data on the touch screen to download the RCC.zip folder containing all RCC data files to the memory stick.

Analyze data

12. Using Excel, import the individual RCC data files into the RCC collector file template provided on the memory stick. Open the RCC collector file in Excel, enable the use of macros by selecting the "Enable this content" radio button, and then click the "Import RCC Files" button.

    All RCC data files that need be analyzed together should be present in the same folder location to multi-select the files and have the same RLF associated with them.

13. If the abundance of target mRNA molecules is going to be compared across multiple different samples, normalize the counts for all target genes in all samples/flow cells based on the positive spike-in controls to account for differences in hybridization efficiency and post-hybridization processing, including purification and immobilization of complexes. Sum the counts for the positive spike-in controls for each sample/flow cell to estimate the overall assay efficiency. Then, calculate a normalization factor for each sample based on the relative number of positive control counts. For each sample/flow cell, multiply the counts for each target gene and the control genes by the normalization factor for that sample/flow cell.
14. Calculate the average normalized background count for a sample/flow cell from the average of the eight negative control counts. Subtract this value from the normalized count of each gene target in that sample/flow cell to yield normalized, background-subtracted counts for each target gene.
15. To determine if the normalized, background-subtracted counts are statistically above background, perform a Student's *t* test against the eight human negative controls.

    A gene is considered to be above background if the average count for the target gene is greater than the average counts for the eight negative control genes and if the P value of the t test is less than 0.05.

The t test assumes that the values being compared follow a normal distribution and that the variances in the data are equal in the two groups being compared. If the data are not normally distributed, the Mann-Whitney U test (also called the Mann-Whitney-Wilcoxon [MWW] or Wilcoxon rank-sum test) can be used, or Welch's t test can be used if the variances between the two groups being compared are highly unequal.

16. Average triplicates for each sample.

17. Calculate relative changes in the abundance of target mRNA molecules using the normalized, background-subtracted counts for one or more reference genes included in the CodeSet.

For an example of RCC data files and data analysis, see the supplementary file at http://www.currentprotocols.com/protocol/mb25B10.

COMMENTARY

Background Information

The development of gene expression signatures that can serve as phenotypes (and therefore experimental read-outs) for various biological states represents a powerful tool for the discovery of functional connections between genes, diseases, and drugs. The underlying assumption is that common gene expression changes reflect functional connectivity, which is expected to reflect a role in regulating common biological processes (Lamb et al., 2003; Amit et al., 2009). Despite the promise and early success of the approach in distinguishing different cancer types and predicting patient survival and response to therapy (van de Vijver et al., 2002; Glas et al., 2005; Burczynski et al., 2006), the robustness of the approach has not yet been evaluated and validated on a large scale. The challenge for the future will be to take these analyses to a higher level by performing high-throughput small-molecule and RNA-interference screens, where measuring the entire transcriptome is neither necessary nor economically viable. An ideal method would provide a direct measurement of each target mRNA of interest, using a small amount of starting material and without requiring cDNA synthesis or amplification. Importantly, the method should be amenable to high-throughput screening with sensitivity that is higher than that of microarrays and about equal to that of TaqMan RT-PCR.

NanoString's nCounter Gene Expression Assay is a novel technology based on direct measurement of target mRNA molecules. The technology uses color-coded molecular barcodes to designate each individual mRNA of interest, and single molecule imaging to detect and count up to 800 unique transcripts in a single reaction. The single-molecule imaging provides a large dynamic range, and the assay as a whole offers significantly higher levels of precision and sensitivity compared to microarray gene expression profiling. Compared to surface-bound hybridization in the case of microarrays, the solution-phase hybridization of this method offers low background signals, significantly increasing the detection of low-abundance mRNAs. Measured expression levels show consistently tight coefficients of variation (CV) across a broad range of relative mRNA abundance levels, ranging from <20% all the way down to a sensitivity of a single copy per cell. One of the biggest advantages of the assay is that it does not require the conversion of mRNA to cDNA or amplification of the resulting cDNA. The assay can be customized to measure target genes defined by the user and offers true multiplexing capability of up to 800 genes in a single reaction. Unlike real-time quantitative PCR, a single reaction is required per sample regardless of the level of multiplexing.

The disadvantage of the nCounter assay is its dependence on specialized equipment and the up-front cost of setting up the technology. The instrumentation (Prep Station and Digital Analyzer) costs approximately $230,000 (U.S.) and would ideally be bought and set up in a core facility rather than individual labs. The price per data point for custom CodeSets decreases as the numbers of targets and samples increase, thus making it economical for querying large sets of targets against large sample sets. For example, the price per data point for an average project (including all necessary reagents) can range from ~$1.00 for a study of 100 genes in 96 samples with a single replicate to approximately $0.30 for a study of 400 genes in 384 samples in triplicate.

The BioMark real-time quantitative PCR system (Fluidigm; *http://www.fluidigm.com*) is a high-throughput system that uses microfluidics and a microfabricated device to perform

thousands of real-time quantitative PCR reactions simultaneously (Golden et al., 2008; Major et al., 2008; Spurgeon et al., 2008). Although the sensitivity achieved by the two platforms is comparable, the Fluidigm system does not allow multiplexing within the same reaction. Alternatively, the FlexMAP detection system (Luminex; *http://www.luminexcorp.com*) offers high-throughput signature gene expression analysis for up to 100 different transcripts across many samples (Landegren et al., 1988; Hsuih et al., 1996; Nilsson et al., 2000). However, this assay has a much lower sensitivity than the nCounter system (more comparable to microarrays) and requires enzymatic conversion of mRNA to cDNA followed by amplification (Peck et al., 2006). In general, Affymetrix GeneChips, the gold standard in microarray technology, can detect a minimum of 10 mRNA copies/cell with good reproducibility (Evans et al., 2002).

There are many applications for this technology, including detection of low-abundance transcripts that cannot be detected by microarrays. Other applications include high-throughput screening (small molecule or RNA-interference) using gene expression signature-based read-outs, chromatin immunoprecipitation assays, splice variation analysis, pathway analysis, fusion transcript expression, and pathogen identification (Amit et al., 2009). In principle, the assay can be used to measure other types of nucleic acids including microRNAs and long non-coding RNAs.

Critical Parameters and Troubleshooting

There are three important parameters in determining the success of this digital, multiplexed gene expression assay. First and foremost is the development of a useful gene expression signature that reflects the biological state of interest. For details on pattern-matching algorithms that are commonly employed to develop robust gene expression signatures from publicly available

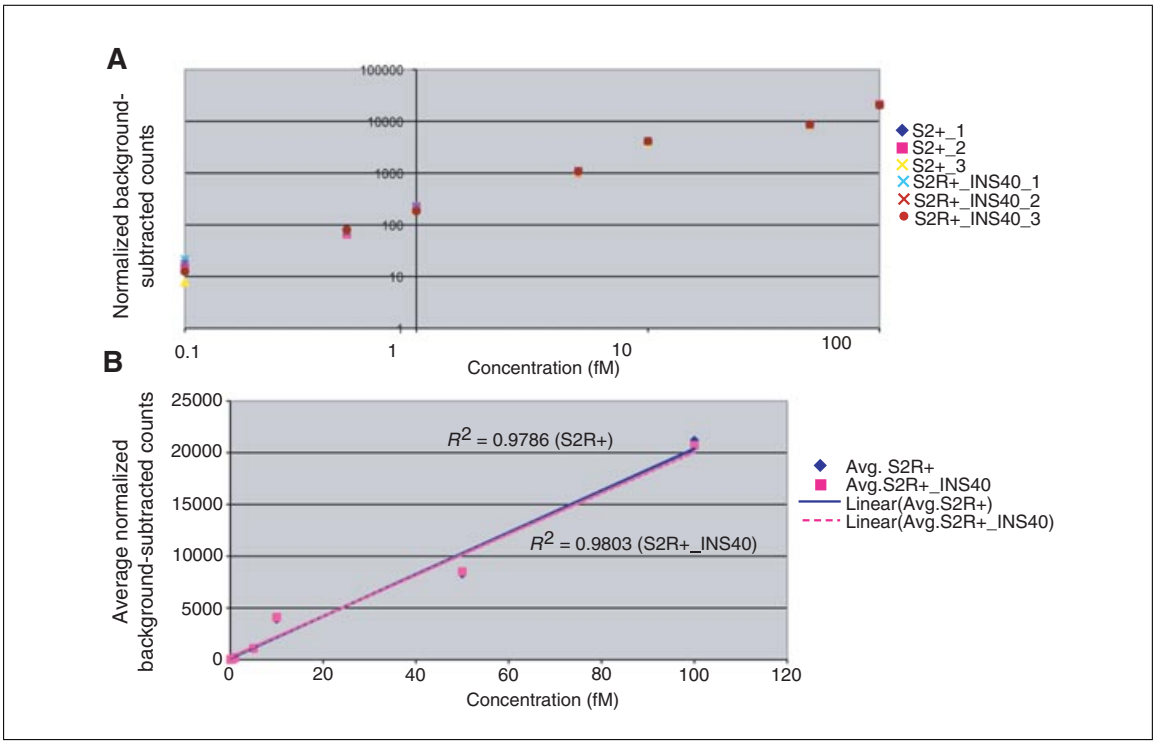

Figure 25B.10.2 Linearity and reproducibility of the positive spike-in controls. DNA oligonucleotide targets were spiked into each sample at concentrations of 0.1, 0.5, 1, 5, 10, 50 and 100 fM. No target was added for the negative control probe pairs. **(A)** Signal (counts) on a log scale vs concentration of the spike on a log scale. Each of three replicate measurements for each spike in baseline *Drosophila* S2R+ cells (S2R+_1, −2, −3) and insulin-stimulated S2R+ cells (S2R+_INS40_1, −2, −3) is shown. The replicate measurements overlap completely except at the lowest concentrations. **(B)** Average signal vs concentration on a linear scale for positive spike-in controls in both baseline and insulin-stimulated *Drosophila* S2R+ samples. The correlation coefficients (R^2) of the linear fit to the average signal are 0.9786 and 0.9803 for baseline and insulin-stimulated samples, respectively. The normalized background-subtracted counts used to construct both graphs are available in the supplementary file at *http://www.currentprotocols.com/protocol/mb25B10*.

microarray data, refer to the review by Nevins and Potti (2007) and the article by Amit et al. (2009). Also refer to public databases such as the Molecular Signatures Database (MSigDB) and Oncomine for collections of available gene sets. The second important parameter is the use of multiple independent probe pairs for each mRNA target of interest and a number of different reference genes that show constant gene expression levels across experiments. Third, as with any gene expression assay, the quality of the total RNA preparation must be verified before proceeding with the hybridization and downstream processing. Results using cell lysate samples are highly correlated with those from total RNA (>0.95) for both low- and high-expressing genes; however, this is only possible if the cell lysate is used immediately after preparation without freeze/thaw cycles. Since this is an enzyme-free system (i.e., no PCR or reverse transcription), it is likely that cell lysate counts correlate better with total RNA in this system than in others.

It is always preferable to have at least three probe pairs per target gene. It is also highly recommended to include multiple (eight to ten) reference genes since it is likely that the level of reference gene mRNA may change under certain conditions. The positive spike control counts can also be used for normalization instead of a reference gene. If data are not reproducible, experiments to be compared should be set up on the same day and run in parallel.

Anticipated Results

The limit of detection for this assay is as low as 0.5 fM based on spike-in controls (roughly corresponding to ~1 copy per cell) and is achieved ~90% of the time (Fig. 25B.10.2 and supplementary file at *http://www.currentprotocols.com/protocol/mb25B10*). The assay offers >1.5-fold change sensitivity for target mRNAs with >5 copies per cell and >2.0-fold change sensitivity for low-abundance mRNAs (1 copy per cell). The assay also has a high degree of reproducibility with $R^2 \geq 0.95$ (Fig. 25B.10.3 and supplementary file) with a linear dynamic range of 7×10^5 total counts.

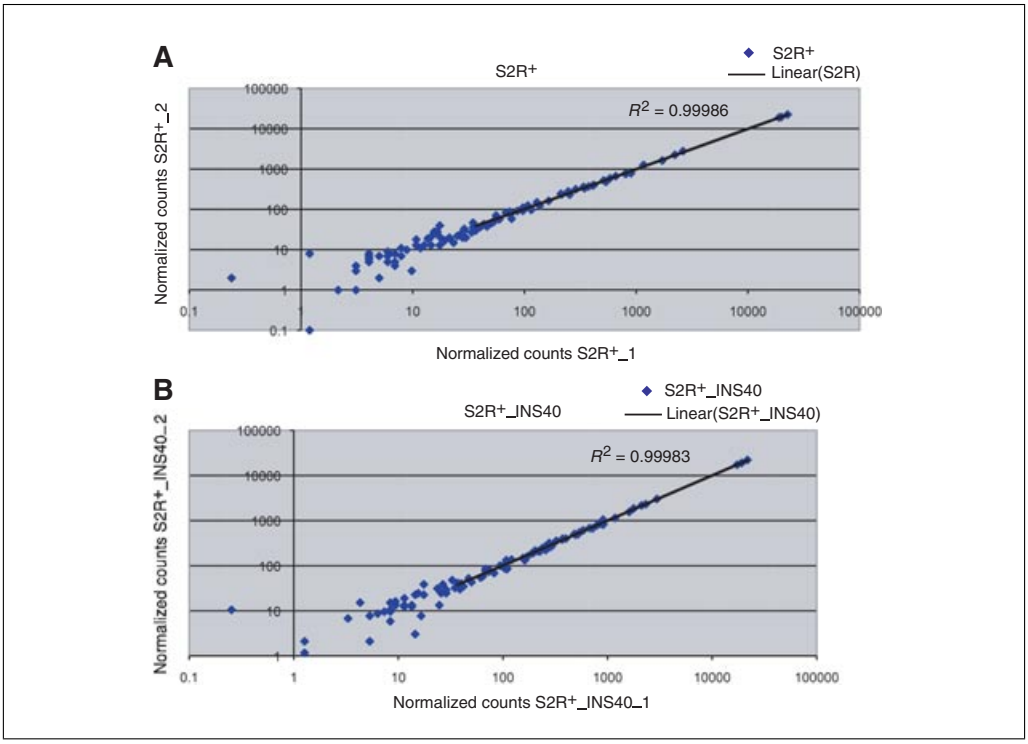

Figure 25B.10.3 Reproducibility of the nCounter platform. **(A)** Scatter plot of normalized background-subtracted counts for all 100 genes assayed, shown in log scale for technical replicates of baseline S2R+ samples (S2R+_1 vs S2R+_2). Genes were not filtered based on present/absent calls. The R^2 value of the linear fit to this data is 0.9999. **(B)** Scatter plot of normalized background-subtracted counts for all 100 genes assayed, shown in log scale for technical replicates of insulin-stimulated S2R+ samples (S2R+_INS40_1 vs S2R+_INS40_2). Genes were not filtered based on detection. The R^2 value of a linear fit to this data is 0.9998. The normalized background-subtracted counts used to construct both graphs are available in the supplementary file at *http://www.currentprotocols.com/protocol/mb25B10*.

Time Considerations

The entire protocol from hybridization to data collection and processing requires 2 days, but hands-on time is less than 30 min. It takes 5-10 min to set up 12 hybridization reactions (excluding the time required to thaw reagents), and the hybridization reaction proceeds overnight for 12-16 hr. The set-up for the Prep Station takes 10 min, and post-hybridization processing of 12 samples requires <2.5 hr. The Digital Analyzer takes 4-5 hr to process 12 samples at a 600 FOV resolution, and can analyze 72 samples per day running unattended in continuous mode. In this case, the hybridizations should be staggered so that when post-hybridization processing of one cartridge is completed on the Prep Station, the next cartridge is ready to be loaded. Processed cartridges can be loaded onto the Digital Analyzer sequentially, even while another cartridge is being imaged.

Literature Cited

Amit, I., Garber, M., Chevrier, N., Leite, A.P., Donner, Y., Eisenhaure, T., Guttmann, M., Grenier, J.K., Li, W., Zuk, O., Schubert, L.A., Birditt, B., Shay, T., Goren, A., Zhang, X., Smith, Z., Deering, R., McDonald, R.C., Cabili, M., Bernstein, B.E., Rinn, J.L., Meissner, A., Root, D.E., Hacohen, N., and Regev, A. 2009. Unbiased reconstruction of a mammalian transcriptional network mediating pathogen responses. *Science* 326:257-263.

Burczynski, M.E., Peterson, R.L., Twine, N.C., Zuberek, K.A., Brodeur, B.J., Casciotti, L., Maganti, V., Reddy, P.S., Strahs, A., Immermann, F., Spinelli, W., Schwertschlag, U., Slager, A.M., Cotreau, M.M., and Dorner, A.J. 2006. Molecular classification of Crohn's disease and ulcerative colitis patients using transcriptional profiles in peripheral blood mononuclear cells. *J. Mol. Diagn.* 8:51-61.

Devonshire, A.S., Elaswarapu, R., and Foy, C.A. 2010. Evaluation of external RNA controls for the standardisation of gene expression biomarker measurements. *BMC Genomics* 11:662.

Evans, S.J., Datson, N.A., Kabbaj, M., Thompson, R.C., Vreugdenhil, E., De Kloet, E.R., Watson, S.J., and Akil, H. 2002. Evaluation of Affymetrix Gene Chip sensitivity in rat hippocampal tissue using SAGE analysis. Serial analysis of gene expression. *Eur. J. Neurosci.* 16:409-413.

Geiss, G.K., Bumgarner, R.E., Birditt, B., Dahl, T., Dowidar, N., Dunaway, D.L., Fell, H.P., Ferree, S., George, R.D., Grogan, T., James, J.J., Maysuria, M., Mitton, J.D., Oliveri, P., Osborn, J.L., Peng, T., Ratcliffe, A.L., Webster, P.J., Davidson, E.H., Hood, L., and Dimitrov, K. 2008. Direct multiplexed measurement of gene expression with color-coded probe pairs. *Nat. Biotechnol.* 26:317-325.

Glas, A.M., Kersten, M.J., Delahaye, L.J., Witteveen, A.T., Kibbelaar, R.E., Velds, A., Wessels, L.F., Joosten, P., Kerkhoven, R.M., Bernards, R., van Krieken, J.H., Kluin, P.M., van't Veer, L.J., and de Jong, D. 2005. Gene expression profiling in follicular lymphoma to assess clinical aggressiveness and to guide the choice of treatment. *Blood* 105:301-307.

Golden, T.R., Hubbard, A., Dando, C., Herren, M.A., and Melov, S. 2008. Age-related behaviors have distinct transcriptional profiles in *Caenorhabditis elegans*. *Aging Cell* 7:850-865.

Hsuih, T.C., Park, Y.N., Zaretsky, C., Wu, F., Tyagi, S., Kramer, F.R., Sperling, R., and Zhang, D.Y. 1996. Novel, ligation-dependent PCR assay for detection of hepatitis C in serum. *J. Clin. Microbiol.* 34:501-507.

Lamb, J., Ramaswamy, S., Ford, H.L., Contreras, B., Martinez, R.V., Kittrell, F.S., Zahnow, C.A., Patterson, N., Golub, T.R., and Ewen, M.E. 2003. A mechanism of cyclin D1 action encoded in the patterns of gene expression in human cancer. *Cell* 114:323-334.

Lamb, J., Crawford, E.D., Peck, D., Modell, J.W., Blat, I.C., Wrobel, M.J., Lerner, J., Brunet, J.P., Subramanian, A., Ross, K.N., Reich, M., Hieronymus, H., Wei, G., Armstrong, S.A., Haggarty, S.J., Clemons, P.A., Wei, R., Carr, S.A., Lander, E.S., and Golub, T.R. 2006. The Connectivity Map: Using gene-expression signatures to connect small molecules, genes, and disease. *Science* 313:1929-1935.

Landegren, U., Kaiser, R., Sanders, J., and Hood, L. 1988. A ligase-mediated gene detection technique. *Science* 241:1077-1080.

Major, M.B., Roberts, B.S., Berndt, J.D., Marine, S., Anastas, J., Chung, N., Ferrer, M., Yi, X., Stoick-Cooper, C.L., von Haller, P.D., Kategaya, L., Chien, A., Angers, S., MacCoss, M., Cleary, M.A., Arthur, W.T., and Moon, R.T. 2008. New regulators of Wnt/beta-catenin signaling revealed by integrative molecular screening. *Sci. Signal.* 11:ra12.

Nevins, J.R. and Potti, A. 2007. Mining gene expression profiles: Expression signatures as cancer phenotypes. *Nat. Rev. Genet.* 8:601-609.

Nilsson, M., Barbany, G., Antson, D.O., Gertow, K., and Landegren, U. 2000. Enhanced detection and distinction of RNA by enzymatic probe ligation. *Nat. Biotechnol.* 18:791-793.

Peck, D., Crawford, E.D., Ross, K.N., Stegmaier, K., Golub, T.R., and Lamb, J. 2006. A method for high-throughput gene expression signature analysis. *Genome Biol.* 7:R61.

Rhodes, D.R. and Chinnaiyan, A.M. 2005. Integrative analysis of the cancer transcriptome. *Nat. Genet.* 37:S31-S37.

Rhodes, D.R., Yu, J., Shanker, K., Deshpande, N., Varambally, R., Ghosh, D., Barrette, T., Pandey, A., and Chinnaiyan, A.M. 2004. Large-scale meta-analysis of cancer microarray data identifies common transcriptional profiles of neoplastic transformation and progression. *Proc. Natl. Acad. Sci. U.S.A.* 101:9309-9314.

Rhodes, D.R., Kalyana-Sundaram, S., Mahavisno, V., Barrette, T.R., Ghosh, D., and Chinnaiyan, A.M. 2005. Mining for regulatory programs in the cancer transcriptome. *Nat. Genet.* 37:579-583.

Spurgeon, S.L., Jones, R.C., and Ramakrishnan, R. 2008. High-throughput gene expression measurement with real time PCR in a microfluidic dynamic array. *PLoSONE* 3:e1662.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., and Mesirov, J.P. 2005. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U.S.A.* 102:15545-15550.

van de Vijver, M.J., He, Y.D., van't Veer, L.J., Dai, H., Hart, A.A., Voskuil, D.W., Schreiber, G.J., Peterse, J.L., Roberts, C., Marton, M.J., Parrish, M., Atsma, D., Witteveen, A., Glas, A., Delahaye, L., van der Velde, T., Bartelink, H., Rodenhuis, S., Rutgers, E.T., Friend, S.H., and Bernards, R. 2002. A gene-expression signature as a predictor of survival in breast cancer. *N. Engl. J. Med.* 347:1999-2009.

Internet Resources

http://www.nanostring.com/

NanoString Technologies website for detailed information regarding products, related literature, and applications.

http://www.broadinstitute.org/gsea/msigdb/index.jsp

The Molecular Signatures Database (MSigDB) is a collection of gene sets, including positional gene sets, curated gene sets, motif gene sets, computational gene sets, and gene ontology gene sets.

https://www.oncomine.org/resource/login.html

A database of cancer gene expression profiles.

http://www.luminexcorp.com

Supports a bead-based platform for high-throughput gene expression signature analysis for the measurement of up to 100 transcripts in many thousands of samples.