Journal of Biological Methods | **2020** | **Vol**. 7(1) | e123
DOI: 10.14440/jbm.2020.300

BENCHMARK

Comparing NGS and NanoString platforms in peripheral blood mononuclear cell transcriptome profiling for advanced heart failure biomarker development

Galyna Bondar[1,2], Wenjie Xu[3], David Elashoff[1], Xinmin Li[1], Emmanuelle Faure-Kumar[1], Tra-Mi Bao[1,2], Tristan Grogan[1], Jim Moose[2], Mario C. Deng[1,2]*

[1]David Geffen School of Medicine, University of California Los Angeles Medical Center, Los Angeles, CA 90095, USA
[2]LeukoLifeDx, Point Pleasant, NJ 08742, USA
[3]Nanostring Technologies, Seattle, WA 98109, USA

*Corresponding author: Mario C. Deng, Email: mdeng@mednet.ucla.edu

Competing interests: The authors Galyna Bondar and Mario C. Deng are co-founding equity holders of LeukoLifeDx, Inc., the developer of MyLeukoMAP™ biomarker test conceptualized in this manuscript. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

Abbreviations used: AdHF, advanced heart failure; CPT, cell preparation tubes; CSV, comma-separated value; DEGs, differentially expressed genes; GEP, gene expression profile; HF, heart failure; HFC, heart failure control; HTx, heart transplant; MCS, mechanical circulatory support; NGS, next generation sequencing; PBMC, peripheral blood mononuclear cell; PBS, phosphate buffer saline; TPM, transcripts per million

Received April 11, 2019; Revision received October 18, 2019; Accepted October 21, 2019; Published January 3, 2020

ABSTRACT

In preparation to create a clinical assay that predicts 1-year survival status of advanced heart failure (AdHF) patients before surgical/interventional therapies and to select the appropriate clinical assay platform for the future assay, we compared the properties of next generation sequencing (NGS) used in the gene discovery phase to the NanoString platform used in the clinical assay development phase. In 25 AdHF patients in a tertiary academic medical center from 2015 to 2016, PBMC samples were collected and aliquoted for NGS RNA whole transcriptome sequencing and compared to 770 genes represented on NanoString's PanCancer IO 360 Gene Expression research panel. Prior to statistical analysis, NanoString and NGS expression values were log transformed. We computed Pearson correlation coefficients for each sample, comparing gene expression values between NanoString and NGS across the set of matched genes and for each of the matched genes across the set of samples. Genes were grouped by average NGS expression, and the NanoString-NGS correlation for each group was computed. Out of 770 genes from the NanoString panel, 734 overlapped between both platforms and showed high intrasample correlation. Within an individual sample, there was an expression-level dependent correlation between both platforms. The low- *vs*. intermediate/high-expression groups showed NGS average correlation 0.21 *vs*. 0.58–0.68, respectively, and NanoString average correlation 0.07–0.34 *vs*. 0.59–0.70, respectively. NanoString demonstrated high reproducibility (R^2 > 0.99 for 100 ng input), sensitivity (probe counts between 100 and 500 detected and quantified), and robustness (similar gene signature scores across different RNA input concentrations, cartridges, and outcomes). Data from NGS and NanoString were highly correlated. These platforms play a meaningful, complementary role in the biomarker development process.

Keywords: advanced heart failure, biomarker, NanoString, next generation sequencing, prediction test

INTRODUCTION

The explosion of high throughput technologies available for generating large-scale molecular measurements has accelerated biomarker development [1]. A biomarker is defined as a biological characteristic that is objectively measured and evaluated as an indicator of biological processes or a pharmacologic response to a therapeutic intervention [2]. An important use of biomarkers is the identification of predictive and prognostic factors in disease management. We are interested in identifying biomarkers in heart failure (HF), and more specifically,

How to cite this article: Bondar G, Xu W, Elashoff D, Li X, Faure-Kumar E, Bao TM, Grogan T, Moose J, Deng MC. Comparing NGS and nanostring platforms in peripheral blood mononuclear cell transcriptome profiling for advanced heart failure biomarker development. *J Biol Methods* 2020;7(1):e123. DOI: 10.14440/jbm.2020.300

advanced heart failure (AdHF).

HF is a cardiovascular syndrome that results from a mismatch between demand and supply of oxygenated blood, leading to progressive organ dysfunction and organ failure [3]. This disease affects more than 20 million people worldwide—6 million persons in the US alone—and is a major public health concern due to its tremendous societal and economic burden. The estimated cost of treating HF was $37 billion in 2009 and is expected to increase to $97 billion by 2030 [4,5].

The mortality risk in AdHF patients depends on a complex combination of demographic, biological and physiological parameters. Some patients with AdHF are, despite having similar HF severity and stage [4,5], more frail [6], immunologically compromised, and susceptible to multi-organ dysfunction. The chronic immune system activation that characterizes and exacerbates AdHF can be measured by gene activity [7-13].

Our previous study shows that, in AdHF patients undergoing mechanical circulatory support implantation, preoperative peripheral blood mononuclear cell (PBMC) gene expression profile (GEP) can assist in predicting early changes in organ function scores and correlates with long-term outcomes. The results indicate that a set of 28 differentially expressed genes (DEGs) can predict day 8 organ function, 105 DEGs correlate to 1-year survival status, and 12 genes overlap between the two gene sets [14].

Now, we aim to create a clinical test that predicts 1-year survival status of individual AdHF patients before any type of HF surgical/interventional therapy, such as mechanical circulatory support surgery or heart transplantation [15-17]. Our HF-survival prediction development strategy consists of a candidate gene discovery phase, followed by a lockdown of a gene list on a clinical-commercial platform.

Gene discovery of the whole-transcriptome using next generation sequencing (NGS) is a key phase of the biomarker test development. NGS is rapidly becoming the method of choice for transcriptional profiling experiments. High throughput sequencing allows identification of novel transcripts and does not require a sequenced genome. Whole-transcriptome analysis can provide information on rare transcripts, splice variants and non-coding RNAs, which can characterize complex phenotypes. NGS can also perform unsupervised RNA GEP tests, which can provide great flexibility, sensitivity, and accuracy in gene expression measurements [18]. In the commercial test development process, NGS is only used to discover candidate genes.

NanoString Technologies is a leading platform for the translation of test development into clinical practice [19]. This closed platform hybridization-system has clearly defined transcript detection and measurements that are not suitable for gene discovery. NanoString has in-built properties that help to avoid bias, because it does require neither library construction, enzymes, nor processing. The NanoString method works in less-than-ideal conditions, because it does not require the conversion of mRNA to cDNA by reverse transcription nor the amplification of the resulting cDNA by PCR. Instead, it is based on direct digital detection of mRNA molecules of interest using target-specific, color-coded probe pairs [18].

The goal in transcriptome biomarker development is an optimal alignment of performance of the discovery and commercial platforms. In the era of DNA-hybridization arrays and RT-PCR used for discovery and commercialization, the concordance was in the range of < 50% of candidate genes [20]. We were interested in examining this question for the state-of-the art contemporary platforms for discovery and commercialization, NGS and NanoString. Therefore, in this study, we compare the general performance of the NGS discovery platform and the NanoString commercial platform, using NGS RNA whole transcriptome sequencing in comparison to 770 genes represented on NanoString's PanCancer IO 360 Gene Expression research panel. Based on the work presented in this paper and further biomarker test development for the identification of DEGs between patients who did and did not survive 1-year post-HF surgical / interventional therapies, subsequent analysis will be needed to confirm the results generated in this study.

MATERIALS & METHODS

Patients

We conducted a study with 25 AdHF patients undergoing guidelines directed medical therapy (n = 11), mechanical circulatory support surgery (n = 4) or heart transplant surgery (n = 10) [20] at UCLA Medical Center between August 2015 and 2016 under UCLA Medical Institutional Review Board-approved protocol number 12-000351. All patient treatments were optimized based on the recommendations of the multidisciplinary heart transplant selection committee. Written informed consent was obtained from each participant [21]. Demographic variables are presented in **Table 1**.

Table 1. Demographics (n = 25).

Patient #	Age (yr)	HF intervention	Gender	1-year survival
1	69	HFC	Male	Yes
2	67	MCS	Male	No
3	73	HFC	Male	No
4	23	HTx	Male	Yes
5	65	HTx	Male	Yes
6	45	HFC	Male	Yes
7	56	HTx	Male	Yes
8	48	HTx	Male	Yes
9	73	HFC	Male	Yes
10	71	HTx	Male	Yes
11	47	HFC	Male	Yes
12	46	HFC	Male	Yes
13	42	HFC	Male	Yes
14	57	HTx	Male	No
15	22	HTx	Male	No
16	21	HFC	Female	Yes
17	59	HTx	Male	Yes
18	57	MCS	Male	Yes
19	66	HTx	Male	Yes
20	60	HFC	Male	Yes
21	53	MCS	Male	Yes
22	50	HTx	Male	Yes
23	64	HFC	Male	Yes
24	21	MCS	Male	Yes
25	69	HTx	Male	Yes

Sample collection, processing and RNA purification

Eight ml of blood was drawn into Vacutainer cell preparation tubes (CPT) (Becton Dickinson, Franklin Lakes, NJ) for purified RNA analysis. PBMC from each sample was purified within 2 h of phlebotomy. We focused on the mixed PBMC population, based on our successful Allomap™ biomarker test development experience [18,20-21].

The collected blood from CPT tubes was mixed and centrifuged at room temperature (22°C) for 20 min at 3000 RPM. The cell layer was collected, transferred to 15 ml conical tubes, re-suspended in cold phosphate buffer saline (PBS) (Sigma-Aldrich, St. Louis, MO) and centrifuged for 20 min at 1135 RPM at 4°C. The cell pellet was re-suspended in cold PBS, transferred into an Eppendorf tube and centrifuged for 20 min at 5.6 RPM at 4°C. The pellet was re-suspended in 0.5 ml RNA Protect Cell Reagent (Qiagen, Valencia, CA) and frozen at −80°C. Then, RNA was isolated from the PBMC using RNeasy Mini Kit (Qiagen, Valencia, CA). The quality of the total RNA was assessed using NanoDrop® ND-1000 spectrophotometer (NanoDrop Technologies, Wilmington, DE) and the concentration using Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA). The samples with concentration above 50 ng/µl, purity 260/280 ~2.0, integrity RIN > 9.0 and average > 9.5 were used in the study. Technical details on this protocol have been published on the LifeSciences Protocol Repository Website at: https://www.protocols.io/ (DOI number dx.doi.org/10.17504/protocols.io.jujcnun). In order to conduct a comparison analysis between the NGS and the NanoString platforms, 25 purified RNA samples were aliquoted in two sets.

PBMC transcriptome next generation sequencing analysis

One set of aliquoted RNA was processed using NGS transcriptome analysis at the UCLA Technology Center for Genomics & Bioinformatics. The NGS mRNA library was prepared with Universal Plus mRNA-Seq kit according to the manufacturer's instructions (NuGen, Redwood City, CA), and 100 ng input material was used. Library construction consists of random fragmentation of the poly A mRNA, followed by cDNA production using random polymers. The cDNA libraries were quantitated using Qubit and size distribution was checked on Bioanalyzer 2100 (Agilent Technologies, Palo Alto, CA). The library was sequenced on HiSeq 2500. Clusters were generated to yield approximately 725 K–825 K clusters/mm^2. Cluster density and quality were determined during the run after the first base addition parameters were assessed. We performed single end sequencing runs to align the cDNA sequences to the reference genome. Generated FASTQ files were transferred to the AdHF Research Data Center where Avadis NGS 1.5 (Agilent, Palo Alto, CA and Strand Scientific, CA) was used to align the raw RNA-Seq FASTQ reads to the reference genome. Reference genome Human hg19 and transcript annotation (gtf file) from UCSC hg19 version 2014-06-02-13-47-56 were used for data normalization.

PBMC transcriptome NanoString nCounter analysis

A second set of aliquoted RNA samples was used for NanoString comparison analysis. Each target gene of interest was detected using a pair of reporter and capture probes that together target a continuous 100 nucleotide sequence. Hybridization between target mRNA and reporter-capture probe pairs was performed at 65°C for 20 h using CT1000 Touch Thermal Cycler (Bio-Rad, CA) according to manufacturer protocol. Post hybridization processing was carried out on a fully automated nCounter Prep station liquid-handling robot. Excess probes were removed and the probe/target complexes were aligned and immobilized in the nCounter cartridge, which was then placed in a digital analyzer for image acquisition and data processing (nCounter Digital Analyzer) as per the manufacturer's protocol. The expression level of a gene was measured by counting the number of times the specific barcode for that gene was detected, and the barcode counts were then tabulated in a comma-separated value (CSV) format. The raw digital count of expression was exported from nSolver v3.0 software for downstream analysis.

CodeSet choice

We chose the NanoString PanCancer IO360 Gene Expression panel (File S1: IO360 gene list and signature descriptions) for two main benefits: first, the IO360 panel has 30+ clinically relevant gene signatures (*e.g.*, TIS signature for anti-PD1 treatment) and immune cell type groups (*e.g.*, T-cell and B-cell signatures) built-in. These signatures can reveal biologically relevant information beyond the single gene level. Second, the IO360 panel was designed and structured similarly to a laboratory-developed test or in-vitro-Diagnostic Multivariate Index Assay. With a synthetic panel standard as the reference sample, IO360 can generate single sample assay reports. The synthetic panel standard can also be used for lot-to-lot calibration.

Experimental design

Cartridge #1 and #2 (24 assays) were used to analyze 24 RNA samples at 100 ng input level (20 ng × 5 µl). Cartridge #3 (12 assays) was used to analyze 2 RNA samples at various dilutions (5–80 ng × 5 µl) of 25–400 ng input level. To assess technical reproducibility, we replicated 100 ng input level (20 ng × 5 µl) of sample #1 and #2 in cartridge #3. In order to check reproducibility across the cartridges, sample #2 (20 ng × 5 µl) was loaded in cartridge #1 and #3. To compare RNA to WB gene expression, one additional cartridge (12 assays) was used to test experimental conditions for WB extracts (**Table 2**).

Statistical analysis

Prior to statistical analysis, NanoString and NGS expression values were log transformed. We computed Pearson correlation coefficients for each sample, comparing gene expression values between NGS and NanoString across the set of matched genes. Gene names (Entrez ID) were used to identify matches in the NGS and NanoString datasets. Next, we computed Person correlations for each of the matched genes across the set of samples. Sets of genes were grouped based upon average expression rank for NanoString across the dataset and the mean NanoString-NGS correlation for each set was computed. Similarly, genes were grouped by average NGS expression and the NanoString-NGS correlation for each group was computed.

RESULTS

Number of overlapping genes for comparison between NanoString and NGS data

Using NanoString panel IO360, out of 750 genes, 734 were matched to NGS genes.

Nanostring/NGS-intrasample correlation

We focused on the correlation between two platforms within one

sample. On average, the correlation was high within the individual 25 samples. The average correlation was 0.904, the minimum correlation is 0.86 and the maximum correlation was 0.92 across the 25 samples used in this study. There was a high correlation of 734 genes in a single sample tested by both assays. The example of this correlation in sample 1 (R^2= 0.8165) is shown by the scatter plot (**Fig. 1**). It is important to note that the true cut-off is not labelled as zero in **Figure 1** due to data transformation. The lower-bound for NGS is at a value of 1, which is an offset placed during transformation and represents genes that are zero in all the samples. The dynamic range of NGS (single digits to hundreds or more) differs from that of NanoString (20 counts to 20 k). However, units do not impact the results of correlation analysis.

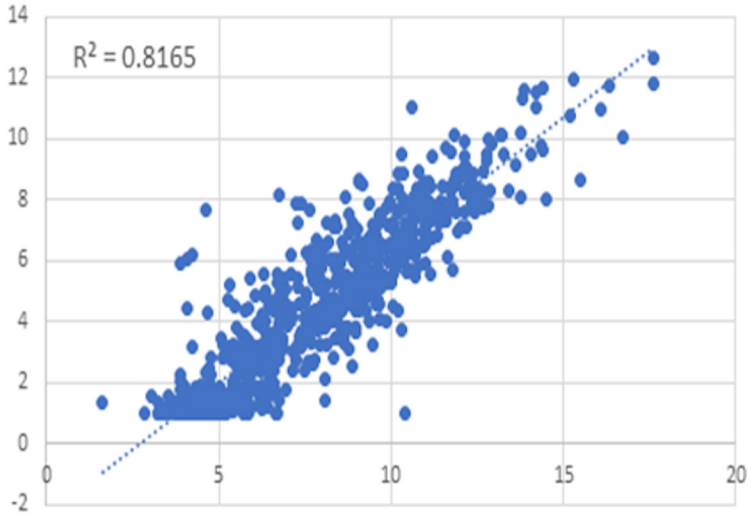

Figure 1. Intrasample correlation of sample 1. X-axis = NanoString, Y-axis = NGS. Correlations were computed for sample 1 to compare gene expression values between NGS and NanoString across 734 overlapping genes.

Table 2. Experimental design of aliquotted PBMC samples.

Lane	Cartridge #1	Cartridge #2	Cartridge #3
1	Sample 2, 20 ng × 5 µl	Sample 14, 20 ng x 5 µl	Sample 1, 5 ng × 5 µl
2	Sample 3, 20 ng × 5 µl	Sample 15, 20 ng x 5 µl	Sample 1, 10 ng × 5 µl
3	Sample 4, 20 ng × 5 µl	Sample 16, 20 ng x 5 µl	Sample 1, 20 ng × 5 µl
4	Sample 5, 20 ng × 5 µl	Sample 17, 20 ng x 5 µl	Sample 1, 20 ng × 5 µl
5	Sample 6, 20 ng × 5 µl	Sample 18, 20 ng x 5 µl	Sample 1, 40 ng × 5 µl
6	Sample 7, 20 ng × 5 µl	Sample 19, 20 ng x 5 µl	Sample 1, 80 ng × 5 µl
7	Sample 8, 20 ng × 5 µl	Sample 20, 20 ng x 5 µl	Sample 2, 5 ng × 5 µl
8	Sample 9, 20 ng × 5 µl	Sample 21, 20 ng x 5 µl	Sample 2, 10 ng × 5 µl
9	Sample 10, 20 ng × 5 µl	Sample 22, 20 ng x 5 µl	Sample 2, 20 ng × 5 µl
10	Sample 11, 20 ng × 5 µl	Sample 23, 20 ng × 5 µl	Sample 2, 20 ng × 5 µl
11	Sample 12, 20 ng × 5 µl	Sample 24, 20 ng × 5 µl	Sample 2, 40 ng × 5 µl
12	Sample 13, 20 ng × 5 µl	Sample 25, 20 ng × 5 µl	Panel standard

Nanostring/NGS-intersample correlation

Grouping definition by gene expression levels: We ranked 734 genes based on average expression levels from low to high. Genes were put into three bins. By NanoString criteria, 128 genes showed low expression level (0–25 counts), 152 genes showed intermediate expression level (26–100 counts) and 454 genes showed high expression level (> 100 counts). **Table 3** summarizes mean expression levels and shows that counts and expression levels are proportional in both, NanoString and NGS platforms.

Nanostring/NGS gene-specific intersample correlation: We examined the gene-specific inter-sample correlation across both platforms. Individual sample plots of gene expression correlation between NGS and NanoString showed the same stratification by expression level. The average correlation was 0.56 across the 734 genes used in this study. The minimum correlation was −0.59 and the maximum correlation was 0.999. **Figure 2** plots the log of the NanoString average expression *vs*.

the correlation between NGS and NanoString. The NanoString results showed that the average correlation between platforms varies by expression level, from low (1–100 counts) $R^2 = 0.07$ to high (> 500 counts) $R^2 = 0.75$ (**Fig. 2A**). The average NGS correlation between platforms varied by log of expression level, from low (< 1 log counts) $R^2 = 0.21$ to high (> 3 log counts) $R^2 = 0.68$ (**Fig. 2B**).

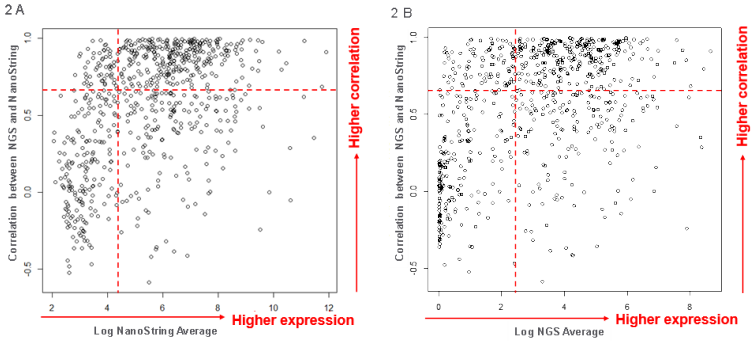

Figure 2. Expression-level dependent correlation for NanoString (A) and NGS (B) perspectives. Gene-specific intersample correlations were computed across NanoString and NGS. Dashed red lines visually separate high and low expressors.

Table 3. Mean gene expression levels in NanoString and NGS platforms.

Number of counts	Number of genes	NanoString mean counts	NGS mean TPM
0–25	128	16.9	1.25
26–100	152	51.1	7.87
> 100	454	2995.8	214.37

The average correlation between NGS and NanoString depended on expression, sorted into NGS and NanoString bins (**Table 4** and **Table 5**). As common practice, the genes were binned in sets of 100 sequentially, from lowest to highest expression. According to NGS log expression average (< 1, 1–2, 2–3, and > 3), one-third of genes in the low expression group showed low correlation between assays, whereas the two-thirds of genes in the intermediate and high-expression groups showed sufficient correlation (**Table 4**). According to the NanoString log expression average (0–3.1, 3.1–3.8, 3.8–4.8, 4.8–5.7, 5.7–6.4, 6.4–7.2, and > 7.2), the average correlation between NGS and NanoString also depended on expression (**Table 5**). The data from both platforms showed that low-expression genes did not yield a high correlation between both assays, whereas high-expression genes did.

Assess assay sensitivity, reproducibility and robustness within the NanoString platform

In order to assess the feasibility of using NanoString for commercial assay development, we demonstrated the sensitivity, reproducibility, and robustness of this platform.

Technical replicates and reproducibility: Technical replicates had R^2 value > 0.98. Equal concentrations (5, 10 and 20 ng × 5 µl) of sample #1 and #2 were used to test the reproducibility of technical replicates. The result showed that the binding density and raw counts increase with the amount of input. All three input levels gave acceptable readout, while 100 ng input gave the most robust signal. The same concentration of 100 ng input level (20 ng × 5 µl) of sample #1 and #2 in cartridge #3 showed technical replicates with the same amount of input and perfect correlation ($R^2 > 0.99$) for both pairs, thus demonstrating extremely high reproducibility (**Fig. 3**).

For sample #1 and #2 in cartridge #3 concentration of 200 ng input level (40 ng × 5 µl), the raw counts increased marginally compared to that of 100 ng input (data not shown). This suggested that given the RNA quantity of samples #1 and #2, 100 ng of input yielded the most precise result and generated robust signal for > 80% genes in the panel without risk of saturation. For this reason, we used 100 ng input for the remainder of this project. The platform was overloaded with 400 ng input (80 ng × 5 µl) (data not shown). The binding density exceeds the limit (2.25) and fields of view dropped due to imaging failure. Raw counts also dropped due to overlapping probes and failure to resolve the barcode.

Sensitivity and limit of detection on the NanoString platform: The background noise level of the assay is around 20 raw counts. The limit of detection on the NanoString platform was set at 2 standard deviations above the mean negative control probe counts. Probe counts ≤ 50 were not quantifiable and represent the lower limit of detection. Probe counts between 50 and 100 were above the limit of detection but could be noisy. Probe counts > 100 were quantifiable and optimal for downstream analysis. **Figure 4** displays a heat map of raw count data without normalization. 128 probes were below the limit of detection (< 50 counts) and labeled as "flag/prune". The absolute limit of detection of NanoString has been published [22]. The intention of this study was to demonstrate that NanoString was a sufficient platform for our sample type and assay requirements.

Robustness and dynamic range on the NanoString platform: To assess the robustness, we calculated gene signature scores and grouped gene signatures according to biological function (**Fig. 5**). The robustness of the results was consistent across different concentrations of RNA input, cartridges, and survival outcomes.

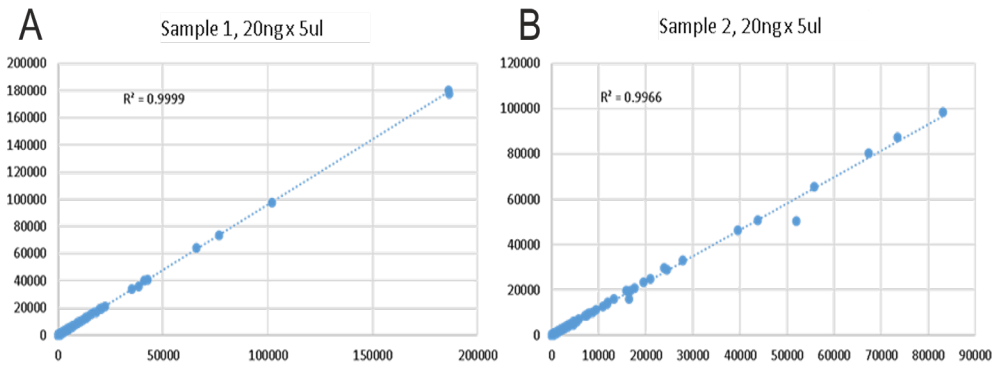

Figure 3. NanoString reproducibility. Equal concentrations of sample #1 (**A**) and #2 (**B**) were used to test the reproducibility of technical replicates.

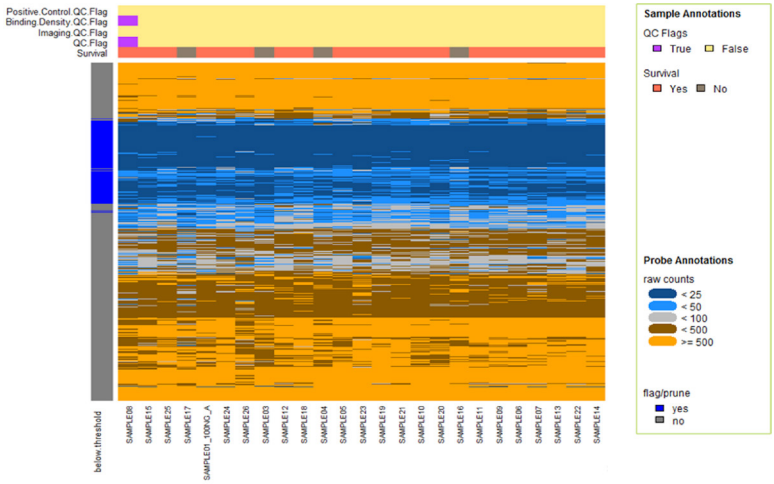

Figure 4. NanoString sensitivity and limit of detection. Raw data form IO360 panel without normalization were shown in a heat map. Probes below the limit of detection (< 50 counts) were labeled as “flag/prune”.

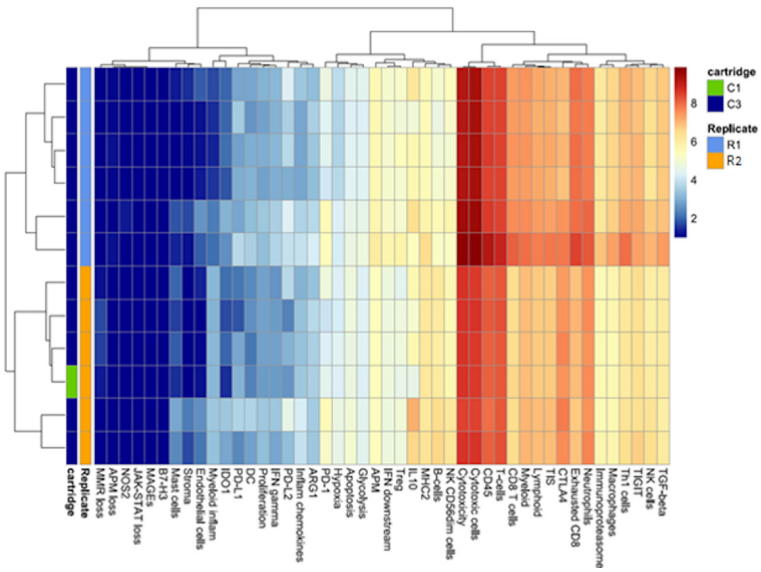

Figure 5. NanoString robustness and dynamic range. Gene signature scores were calculated and grouped according to biological function across different concentrations of RNA input, cartridges, and survival outcomes. C1, cartridge #1; C3, cartridge #3; R1, replicate of sample #1; R2, replicate of sample #2.

Table 4. Average correlation in NGS platform.

Log NGS expression average	Average correlation	Expression level
< 1	0.21	Low
1–2	0.58	Intermediate
2–3	0.63	High
> 3	0.68	High

Table 5. Average correlation in NanoString platform.

Gene order (average expression among 734 genes)	Log nanostring expression average	Average correlation	Expression level
1–100	0–3.1	0.07	Low
100–200	3.1–3.8	0.34	Low
200–300	3.8–4.8	0.59	Intermediate
300–400	4.8–5.7	0.64	High
400–500	5.7–6.4	0.67	High
500–600	6.4–7.2	0.75	High
Above 600	> 7.2	0.70	High

All data were normalized by geometric mean of the built-in housekeeping genes. The heat map shows an unsupervised clustering of outcomes, with sample #1 representing one year survival status and sample #2 representing one year non-survival. The results were robust within the two groups at different RNA concentrations.

Considering the difficulty of performing a controlled degradation, we simulated degradation by diluting high-quality RNA input to reduce the RNA detectability and the effective target number. Cartridge #3 represented samples from the same RNA prep tubes (sample #1 and sample #2), but loaded at different input amounts (25 ng to 400 ng total RNA). All samples of varying dilution levels produced similar gene signature scores, suggesting a large dynamic range and robustness of the NanoString platform.

To show the inter-cartridge robustness, we placed an identical concentration of sample #2 (100 ng) in cartridge #1 and cartridge #3, which showed excellent technical replication.

DISCUSSION

In this paper, we compared the performance of NGS and NanoString in PBMC transcriptome profiling, using NGS RNA whole transcriptome sequencing to compare the 770 genes represented on NanoString's PanCancer IO 360 Gene Expression panel. Our study showed good correlation between NGS and NanoString data. Out of 750 genes from the NanoString panel IO360, 734 overlapped between both platforms and showed high intrasample correlation. Within an individual sample, there was an expression-level dependent correlation between both platforms. Binning the genes into low, intermediate and high expression levels, the two platforms showed variable correlation, with worse correlation in genes with low expression levels and high correlation in genes with high expression levels. The intermediate and high-expression groups showed NGS average correlation from 0.58–0.68 and NanoString average correlation from 0.59–0.70. The low expression groups showed NGS average correlation of 0.21 and NanoString average correlation from 0.07–0.34. For these reasons, we will use only the intermediate- and high-expression groups (approximately 50%–60% of total expression groups) as targets for commercialization.

Consistent with published studies [19,23], we demonstrated that the reproducibility, sensitivity, and robustness of NanoString makes it a suitable platform for commercial assay development. Technical replicates with equal concentration input levels within one cartridge demonstrated extremely high reproducibility. The RNA input of 100 ng yielded the most precise result, generated robust signal for > 80% genes in the panel without risk of saturation and therefore was used to analyze the remaining 23 samples. In future studies, we plan to use 100 ng input to build our commercial assay. We may increase input to 200 ng for samples with partially degraded RNA. Notably, the calculated gene signature scores were highly consistent across a wide range of concentrations of RNA input, consumables and samples with different clinical outcomes.

Our contemporary biomarker test development holds several benefits over older techniques. For example, the first clinically approved molecular expression test, AlloMap, which was initiated in 2001, used microarray and RT-PCR for gene discovery, validation, and commercialization [20]. At that time, the best-available discovery platform was based on hybridization and was only able to measure the expression of 7300 genes. The transition to a commercial platform using RT-qPCR yielded roughly 25%–30% of genes with similar expression. Our current study utilized expression-level dependent correlation, which suggests that both NGS and NanoString platforms yield results that are more precise, robust, and reproducible for gene discovery and commercialization. As a projection for a commercial transcriptome test development, our study demonstrated an excellent cross talk between gene discovery

and commercial platforms.

This paper only compares purified mRNA, not whole blood. We will perform the platform comparison in a separate project using whole blood.

The expression of very highly expressed groups cannot be quantified above 500 raw probe annotation counts by using NanoString.

As the research has demonstrated, NGS and NanoString have complementary roles. NGS for discovery, NanoString for commercialization. These two platforms "talk to each other".

Acknowledgments

The NanoString experiment was performed at the IMT core/Center for Systems Biomedicine, which is supported by CURE/P30 DK041301. Funding for the study was obtained by UCLA NIH R21 1R21HL120040-01 (MCD) (PI Deng), UCLA R01 (PI Weiss, Joint PI Deng), UCLA R01 (PI Ping, Co-I Deng), UCLA DOM Internal Funds and the Advanced HF Research Gift to Columbia University (Philip Geier, John Tocco and Robert Milo) and Advanced HF Research Gift to UCLA (Juan Mulder, Peter Schultz, Larry Layne, James & Candace Moose).

References

1. Ghosh D, Poisson LM (2008) "Omics" data and levels of evidence for biomarker discovery. Genomics 93: 13-16. doi: 10.1016/j.ygeno.2008.07.006. PMID: 18723089
2. Biomarkers Definitions Working Group (2001) Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. Clin Pharmacol Ther 69: 89-95. doi: 10.1067/mcp.2001.113989. PMID: 11240971
3. Doehner W, Frenneaux M, Anker SD (2014) Metabolic impairment in heart failure: the myocardial and systemic perspective. J Am Coll Cardiol 64: 1388-1400. doi: 10.1016/j.jacc.2014.04.083. PMID: 25257642
4. Yancy CW, Jessup M, Bozkurt B, Butler J, Casey Jr DE, et al. (2017) 2017 ACC/AHA/HFSA focused update of the 2013 ACCF/AHA guideline for the management of heart failure: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines and the Heart Failure Society of America. J Card Fail 23: 628-651. doi: 10.1016/j.cardfail.2017.04.014. PMID: 28461259
5. Yancy CW, Jessup M, Bozkurt B, Butler J, Casey Jr DE, et al. (2013) 2013 ACCF/AHA guideline for the management of heart failure: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines. J Am Coll Cardiol 62: doi: 10.1016/j.jacc.2013.05.019. PMID: 23747642
6. Joseph SM, Rich MW (2017) Targeting Frailty in Heart Failure. Curr Treat Options Cardiovasc Med 19: 31. doi: 10.1007/s11936-017-0527-5. PMID: 28357683
7. Levine B, Kalman J, Mayer L, Fillit HM, Packer M (1990) Elevated circulating levels of tumor necrosis factor in severe chronic heart failure. N Engl J Med 323: 236-241. doi: 10.1056/NEJM199007263230405. PMID: 2195340
8. Deng MC, Dasch B, Erren M, Möllhoff T, Scheld HH (1996) Impact of left ventricular dysfunction on cytokines, hemodynamics, and outcome in bypass grafting. Ann Thorac Surg 62: 184-190. doi: 10.1016/0003-4975(96)00230-5. PMID: 8678641
9. Caruso R, Trunfio S, Milazzo F, Campolo J, De Maria R, et al. (2010) Early expression of pro- and anti-inflammatory cytokines in left ventricular assist device recipients with multiple organ failure syndrome. ASAIO J 56: 313-318. doi: 10.1097/MAT.0b013e3181de3049. PMID: 20445439
10. Caruso R, Verde A, Cabiati M, Milazzo F, Boroni C, et al. (2012) Association of pre-operative interleukin-6 levels with interagency registry for mechanically assisted circulatory support profiles and intensive care unit stay in left ventricular assist device patients. J Heart Lung Transplant 31: 625-633. doi: 10.1016/j.healun.2012.02.006. PMID: 22386451
11. Kaur K, Dhingra S, Slezak J, Sharma AK, Bajaj A, et al. (2008) Biology of TNFalpha and IL-10, and their imbalance in heart failure. Heart Fail Rev 14: 113-123. doi: 10.1007/s10741-008-9104-z. PMID: 18712475
12. Soejima H, Irie A, Fukunaga T, Oe Y, Kojima S, et al. (2007) Osteopontin expression of circulating T cells and plasma osteopontin levels are increased in relation to severity of heart failure. Circ J 71: 1879-1884. doi: 10.1253/circj.71.1879. PMID: 18037740
13. Mann DL (2002) Inflammatory mediators and the failing heart: past, present, and the foreseeable future. Circ Res 91: 988-998. doi: 10.1161/01.RES.0000043825.01705.1B. PMID: 12456484
14. Bondar G, Togashi R, Cadeiras M, Schaenman J, Cheng RK, et al. (2017) Association between preoperative peripheral blood mononuclear cell gene expression profiles, early postoperative organ function recovery potential and long-term survival in advanced heart failure patients undergoing mechanical circulatory support. PLoS One 12: doi: 10.1371/journal.pone.0189420. PMID: 29236770
15. Deng MC (2004) Heart transplantation: the increasing challenges of evidence-based decision-making. J Am Coll Cardiol 43: 803-805. doi: 10.1016/j.jacc.2003.12.012. PMID: 14998620
16. Deng MC, De Meester JM, Smits JM, Heinecke J, Scheld HH (2000) Effect of receiving a heart transplant: analysis of a national cohort entered on to a waiting list, stratified by heart failure severity. BMJ 321: 540-545. doi: 10.1136/bmj.321.7260.540. PMID: 10968814
17. Deng MC (2018) A peripheral blood transcriptome biomarker test to diagnose functional recovery potential in advanced heart failure. Biomark Med 12: 619-635. doi: 10.2217/bmm-2018-0097. PMID: 29737882
18. Tachibana C (2015) Transcriptomics today: Microarrays, RNA-seq, and more. Science 349: 544-546. doi: 10.1126/science.349.6247.544.
19. Wallden B, Storhoff J, Nielsen T, Dowidar N, Schaper C, et al. (2015) Development and verification of the PAM50-based Prosigna breast cancer gene signature assay. BMC Med Genomics 8: 54. doi: 10.1186/s12920-015-0129-6. PMID: 26297356
20. Deng MC, Eisen HJ, Mehra MR, Billingham M, Marboe CC, et al. (2006) Noninvasive discrimination of rejection in cardiac allograft recipients using gene expression profiling. Am J Transplant 6: 150-160. doi: 10.1111/j.1600-6143.2005.01175.x. PMID: 16433769
21. Adigopula S, Vivo RP, DePasquale EC, Nsair A, Deng MC (2014) Management of ACCF/AHA Stage C heart failure. Cardiol Clin 32: 73-93. doi: 10.1016/j.ccl.2013.09.012. PMID: 24286580
22. Geiss GK, Bumgarner RE, Birditt B, Dahl T, Dowidar N, et al. (2008) Direct multiplexed measurement of gene expression with color-coded probe pairs. Nat Biotechnol 26: 317-325. doi: 10.1038/nbt1385. PMID: 18278033
23. Veldman-Jones MH, Brant R, Rooney C, Geh C, Emery H, et al. (2015) Evaluating Robustness and Sensitivity of the NanoString Technologies nCounter Platform to Enable Multiplexed Gene Expression Analysis of Clinical Samples. Cancer Res 75: 2587-2593. doi: 10.1158/0008-5472.CAN-15-0262. PMID: 26069246

Supplementary information

File S1. IO360 gene list and signature descriptions.
Supplementary information of this article can be found online at http://www.jbmethods.org/jbm/rt/suppFiles/300.