STAR: Multiplexed Quantitative PCR-Based Platform for Gene Expression Analysis and Molecular Diagnostics**

Primera BIOSYSTEMS

Elizabeth P. Garcia[1], Lori A. Dowding[1], Lawrence W. Stanton[2], Vladimir I. Slepnev[1]
[1]Primera BioSystems, 4 Richmond Square, Providence, Rhode Island, [2]Genome Institute of Singapore, 60 Biopolis Street, Singapore

ABSTRACT

We report the development of a new technology for simultaneous quantitative detection of multiple targets in a single sample. STAR, Scalable Transcriptional Analysis Routine, represents a novel integration of RT-PCR and capillary electrophoresis allowing detection of dozens of gene transcripts in a multiplexed format using amplicon size as an identifier for each target. STAR demonstrated similar or better sensitivity and precision compared to two commonly used methods, SYBR Green-based and Taqman probe-based real-time RT-PCR. STAR is a flexible platform for building a variety of applications to monitor gene expression: from single gene assays to assays analyzing the expression level of multiple genes. Using SARS as a model system, STAR technology detected single copies of the viral genome in a multiplexed assay. In blind studies using RNA extracted from various tissues of a SARS-infected individual, STAR correctly identified all samples containing SARS virus, while yielding negative results for non-SARS control samples. Using alternate priming strategies, STAR technology can be adapted to transcriptional profiling studies without requiring a priori sequence information. STAR technology offers a flexible platform for development of highly multiplexed assays in gene expression analysis and molecular diagnostics.

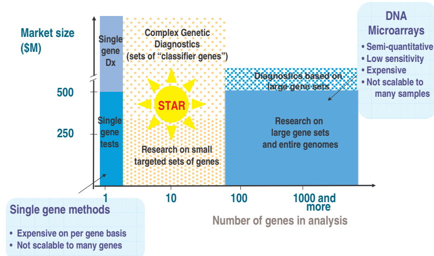
INTRODUCTION

Gene expression analysis is used broadly and extensively as an sensitive and specific diagnostic tool to detect pathogens or specific disease-associated genes in humans. The competitive landscape of available technologies is characterized by clear separation of very sensitive and quantitative methods measuring single bioanalytes (real-time PCR) and methods capable of multiplexing thousands of genes (DNA microarrays) with less sensitivity and quantitative ability compared to single-gene methods. The recent introduction of technologies capable of analyzing a number of analytes or biomarkers in a high-throughput multiplex configuration in clinical diagnostics and gene expression analysis demonstrates a growing trend toward development of multiplex assays. The future of NAT requires better multiplexing abilities that maintain or exceed the current levels of sensitivity.

We have developed STAR (Scalable Transcription Analysis Routine), a gene expression analysis platform that combines the desirable traits of both real-time PCR (precision, sensitivity and quantification) and DNA microarray (miltiplexing) into a single system. STAR represents an innovative integration of real-time multiplex PCR and capillary electrophoresis (CE), allowing the simultaneous quantitative measurement of multiple targets in a single sample with high sensitivity. Specificity of PCR amplification is due to appropriate primer choice and reaction conditions. Because CE allows accurate size determination of fluorescently labeled nucleic acids from 50 to 1000 bases with the precision of one base, assays can be conducted simultaneously for dozens of targets whose identities are defined by the specific size of its corresponding amplicons, while maintaining quantification capabilities equal to or better than those observed with established real-time PCR methods. STAR is fast, cost-effective, and has a large dynamic range. Here we present STAR technology and its application to diagnostics and gene expression analysis.

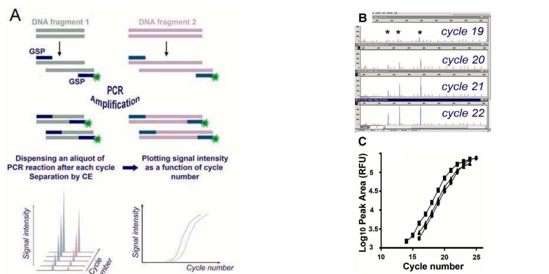
Gene Expression Market Opportunity

Figure 1. Role of STAR Technology In Today's Market



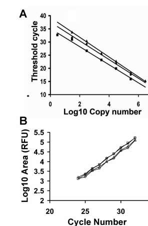
Description of STAR Technology and Proof of Concept

Figure 2. Diagrammatic Representation of STAR Technology



(A) Diagrammatic representation of the STAR process. (B) Electropherograms derived from sequential multiplex PCR cycles. Peaks representing arc, homer1a, and zif268 are marked with an asterisk. Small repeating peaks represent DNA molecular size markers. (C) Amplification curves for arc (closed triangles), homer1a (open circles) and zif268 (closed squares) were reconstructed by plotting the area under each peak against cycle number.

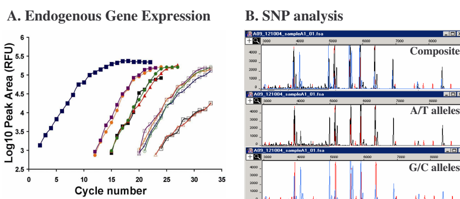
Figure 3. STAR is Sensitive and Reproducible.



(A) Three transcripts were ten-fold serially diluted from 3,000,000 to 3 copies in *E.coli* tRNA (20ug/ml) and amplified as a three-gene multiplex STAR assay. C_T vs. copy number plots are shown for each transcript. Amplification efficiency = $10^{(-1/slope)}-1$. (B) To demonstrate reproducibility, 4 samples, each containing 300 copies of three different transcripts in a background of *E.coli* tRNA (20ug/ml), were multiplex amplified as above. Each amplification curve is represented by open circles, squares, triangles or diamonds.

Multiplexing using STAR Technology

Figure 4. STAR Technology for gene expression and SNP analysis



(A)Twelve endogenous genes of varying expression levels were multiplex amplified using gene specific primers from 400ng total rat brain RNA in a one-step STAR protocol. One of each primer pair was fluorescently labeled. Aliquots from cycles 2 to 33 were analyzed by CE to generate amplification curves. (B) The genotype of 10 different SNPs were monitored using STAR technology using allele-specific primers from 1ng of human genomic DNA pooled from 80 individuals.

Diagnostic Applications of STAR Technology

Figure 5. Detection of SARS Co-V using STAR Technology

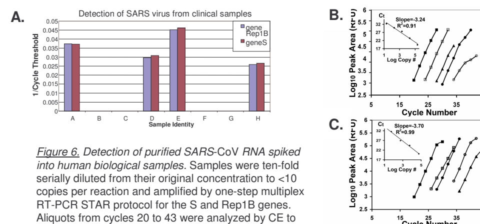
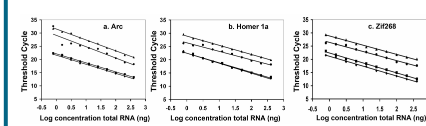


Figure 6. Detection of purified SARS-CoV RNA spiked into human biological samples. Samples were ten-fold serially diluted from their original concentration to <10 copies per reaction and amplified by one-step multiplex RT-PCR STAR protocol for the S and Rep1B genes. Aliquots from cycles 20 to 43 were analyzed by CE to generate amplification curves. Data for the S gene are shown for (A) stool [200,000 {closed squares}; 20,000 {open squares}; 2,000 {closed triangles}; 200 {closed circles}; 20 {open circles} copies] or (B) sputum derived samples [56,000 {closed squares}; 5,600 {open squares}; 560 {closed circles}; 56 {open circles}; 5.6 {closed triangles} copies]. (Insets) C_T vs. copy number graphs derived from data in (A) and (B).

STAR Technology compared to other Real-Time Methods

Figure 6. STAR Technology is Comparable to Taqman probe-based and SYBR Green-based real-time RT-PCR.

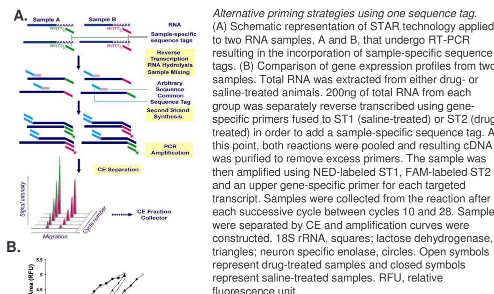


	homer1a			zif268			arc		
	slope	*efficiency*	*r2*	*slope*	*efficiency*	*r2*	*slope*	*efficiency*	*r2*
TaqMan	-3.46 ± 0.05	94.33 ± 1.8	0.998	-3.45 ± 0.15	94.94 ± 5.4	0.996	-4.13 ± 0.10	74.73 ± 2.4	0.995
SYBR	-3.21 ± 0.16	104.8 ± 7.3	0.981	-3.49 ± 0.08	93.56 ± 3.0	0.986	-3.45 ± 0.37	94.88 ± 14	0.878
STAR single	-3.79 ± 0.09	83.41 ± 2.6	0.996	-3.65 ± 0.07	88.07 ± 2.4	0.997	-3.45 ± 0.09	95.10 ± 3.3	0.997
STAR multiplex	-3.61 ± 0.11	89.17 ± 3.8	0.992	-3.73 ± 0.07	85.34 ± 2.3	0.997	-3.43 ± 0.06	95.78 ± 2.4	0.995

Detection of endogenous levels of (A) arc, (B) homer1a and (C) zif268 in rat brain total RNA were assessed by 3 real-time PCR methods: SYBR (circles), Taqman (triangles), and STAR either as an individual reaction (squares) or as part of a multiplex reaction (diamonds). C_Ts were determined from PCR amplifications performed from two-fold serially diluted total rat brain RNA (400 to 0.78ng) and plotted.

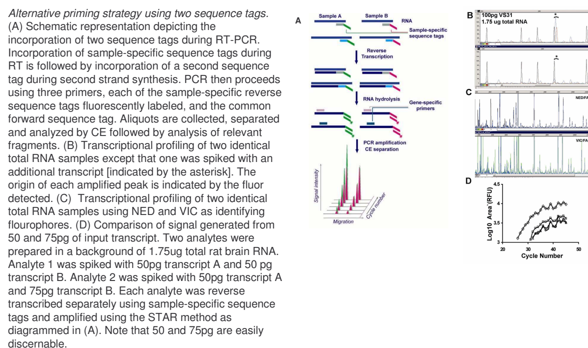
Applications in Gene Expression Profiling

Figure 7. Sample-to-sample Comparison Using STAR Technology



Alternative priming strategies using one sequence tag. (A) Schematic representation of STAR technology applied to two RNA samples, A and B, that undergo RT-PCR resulting in the incorporation of sample-specific sequence tags. (B) Comparison of gene expression profiles from two samples. Total RNA was extracted from either drug- or saline-treated animals. 200ng of total RNA from each group was separately reverse transcribed using gene-specific primers fused to ST1 (saline-treated) or ST2 (drug-treated) in order to add a sample-specific sequence tag. At this point, both reactions were pooled and resulting cDNA was purified to remove excess primers. The sample was then amplified using NED-labeled ST1, FAM-labeled ST2 and an upper gene-specific primer for each targeted transcript. Samples were collected from the reaction after each successive cycle between cycles 10 and 28. Samples were separated by CE and amplification curves were constructed. 18S rRNA, squares; lactose dehydrogenase, triangles; neuron specific enolase, circles. Open symbols represent drug-treated samples and closed symbols represent saline-treated samples. RFU, relative fluorescence unit.

Figure 8. Transcriptional Profiling using STAR Technology



Alternative priming strategy using two sequence tags. (A) Schematic representation depicting the incorporation of two sequence tags during RT-PCR. Incorporation of sample-specific sequence tags during RT is followed by incorporation of a second sequence tag during second strand synthesis. PCR then proceeds using three primers, each of the sample-specific reverse sequence tags fluorescently labeled, and the common forward sequence tag. Aliquots are collected, separated and analyzed by CE followed by analysis of relevant fragments. (B) Transcriptional profiling of two identical total RNA samples except that one was spiked with an additional transcript [indicated by the asterisk]. The origin of each amplified peak is indicated by the fluor detected. (C) Transcriptional profiling of two identical total RNA samples using NED and VIC as identifying flourophores. (D) Comparison of signal generated from 50 and 75pg of input transcript. Two analytes were prepared in a background of 1.75ug total rat brain RNA. Analyte 1 was spiked with 50pg transcript A and 50 pg transcript B. Analyte 2 was spiked with 50pg transcript A and 75pg transcript B. Each analyte was reverse transcribed separately using sample-specific sequence tags and amplified using the STAR method as diagrammed in (A). Note that 50 and 75pg are easily discernable.

CONCLUSIONS

- Improved specificity to allow for more precise identification of a given gene
- Improved sensitivity to allow identification of genes at very low levels
- Scalability to allow analysis for up to one hundred genes in a single experiment with negligible incremental costs for additional analytes or biomarkers
- Rapid development and integration of new gene assays thereby allowing timely expandability to incorporate emerging pathogens, clinical needs, scientific discoveries
- Reduced overall cost relative to other molecular diagnostic platforms

**Significant portions of the work presented here has been accepted for publication in *The Journal of Molecular Diagnostics, 9650 Rockville Pike, Bethesda, Maryland, USA 20814-3993*