APPLICATION INFORMATION

MULTIPLEXED, QUANTITATIVE GENE EXPRESSION ANALYSIS FOR LETTUCE SEED GERMINATION ON GENOMELAB™ GEXP GENETIC ANALYSIS SYSTEM

Eiji Hayashi[1], Natsuyo Aoyama[1], Yong Wu[2], Han-Chang Chi[2], Scott K. Boyer[2], David W. Still[1]
1. California State Polytechnic University, Pomona, Department of Plant Sciences
2. Beckman Coulter, Inc.

Introduction

Gene expression is used to analyze the function of one or more gene(s), determine transcriptional regulation, elucidate signal transduction pathways, map expression-level polymorphisms and aid in the area of molecular medicine, disease diagnosis and treatment. Many traits studied by scientists are physiologically and genetically complex and elucidating the underlying genetic mechanisms is an active area of research. Diabetes, cancer, Alzheimer's and schizophrenia are examples of such complex disorders in which the genes associated with these diseases remain largely unknown. As a first step in the discovery process it is common to employ genome-wide transcript profiling to develop a list of candidate genes.This approach has been used to advance the understanding of complex diseases in humans and has been instrumental in the development of molecular medicine. For example, transcription analysis has been used for clinical purposes, to aid in the diagnoses of cancers and to understand the biology of different treatments (Bigler et al. 2003). Gene expression profiling can be used to classify tumors (Golub et al. 1999; Watson et al. 2001) and in doing so has shown that tumors are far more heterogeneous than previously suspected (Jazaeri et al. 2002). Further, expression profiling of as few as 30 genes has led to the prediction of the biological response of tumors in response to chemotherapy agents (Wang et al. 2002).

Similarly, plant biologists seek to understand the genetic underpinnings of complex physiological processes such as the control of vegetative and reproductive growth, the response of plants to biotic and abiotic stress (Casati et al, 2006), and how to increase the yield of agronomic and horticultural plants. Genome-wide microarrays have successfully identified genes that are differentially expressed under various experimental conditions.Targeted approaches can then be used to test specific hypotheses about function of candidate genes.

Gene expression is most commonly performed by analyzing a single gene at a time using SYBR-green detection. Real-time reverse-transcription quantitative PCR (real-time qPCR) is widely used for gene expression, having replaced Northern blots as the preferred method by which gene expression is quantified. Real-time qPCR has proven to be reproducible, sensitive, and linear over approximately seven orders of magnitude (Wong and Medrano 2005). However, single gene analysis is not practical for medium to high-throughput applications in terms of the amount of time, labor and cost required to process the samples. Because multigenic traits are affected by gene-environment interactions, discovering its biological basis requires processing large number of samples taken from multiple environments. In such cases that require a moderately large number of genes to be assayed across a moderately large number of individuals or samples, a medium to high-throughput method is needed. The GeXP approach is ideally suited for those situations.

In the study presented here, the GenomeLab™ GeXP Genetic Analysis System was used for gene expression analysis of lettuce seed germination. We designed a 15-gene lettuce germination panel and applied gene expression profiling analysis on RNA samples from seed imbibed under red or far-red light, conditions that promote or prevent germination, respectively. The GeXP method was also compared with real-time qPCR in performance, reproducibility, throughput, and cost. Data presented here demonstrate that not only does GeXP accurately and reproducibly perform gene transcript analysis, it is also time and cost-efficient.

Overview of Gene Expression on GeXP

GeXP employs eXpress Profiling (XP-PCR), a patented technology (Figure 1) for multiplex gene expression profiling analysis by which up to 30 genes can be easily multiplexed in the same reaction. There are five basic steps involved: 1) Primer design; 2) cDNA synthesis; 3) PCR; 4) Separation on the GenomeLab GeXP Genetic Analysis System; 5) Fragment Analysis and Expression Profiling Analysis (Figure 2). Detailed instructions for each of these steps are illustrated in the kit insert, and are also accessible online at www.beckmancoulter.com, under Genomics. The following is a brief description of each of these steps.

1) Primer design:
Primers for the multiplexed panel are designed by importing the target gene ID or sequence into the eXpress Designer module of the eXpress Profiler software (Figure 3). The core of the eXpress Designer software is based on Primer3 software algorithms (Rozen and Skaletsky 2000) with a multiplex function added. The primers are designed to generate amplified products with similar GC content and melting temperature. Each primer designed will be appended with a universal primer sequence by the software. These new "chimeric" primers, now each contain the gene-specific sequence and the universal primer sequence. The amplified products are designed to generate gene fragment with lengths between 100-400 nt.

2) cDNA synthesis:
Following RNA extraction, quantification, and normalization, multiplex cDNA synthesis is performed using the reverse-transcriptase (RT) reaction mix containing the gene-specific chimeric reverse primer mix, *KAN*^{r} as an internal control, reverse transcriptase, buffers and template RNA (Table 1). The RT reaction is performed in a thermal-cycler with the following program: 48ºC for 1 min; 37ºC for 5 min; 42ºC for 60 min; 95ºC for 5 min; hold at 4ºC.

3) PCR:
An aliquot (9.3 µL) of the RT reaction is then transferred to the PCR reaction mix which contains $MgCl_2$, the gene-specific forward chimeric primer plex, fluorescently-labeled universal forward primer, unlabeled universal reverse primer and Thermo-Start® DNA polymerase (ABgene) (Table 2). The 96-well plate containing the PCR reaction mixture is transferred to a thermal-cycler and run under the following program: 1 cycle of 95ºC for

Table 1. *RT reaction setup.*

RT Reaction Mix	Volume Per Well
DNase/RNase Free H_2O	3 µL
RT Buffer 5X	4 µL
Custom RT Rev Primer Plex*	2 µL
Reverse Transcriptase	1 µL
KAN^r RNA with RI	5 µL
Sample RNA (5-20 ng/µL)	5 µL
	(25-100 ng total)
Total	**20 µL**

***Note.** The oligo concentration for each un-attenuated reverse primer in the RT Rev Primer Plex is 500 nM. The oligo concentration for the attenuated reverse primer is gene dependent (see Attenuation section).*

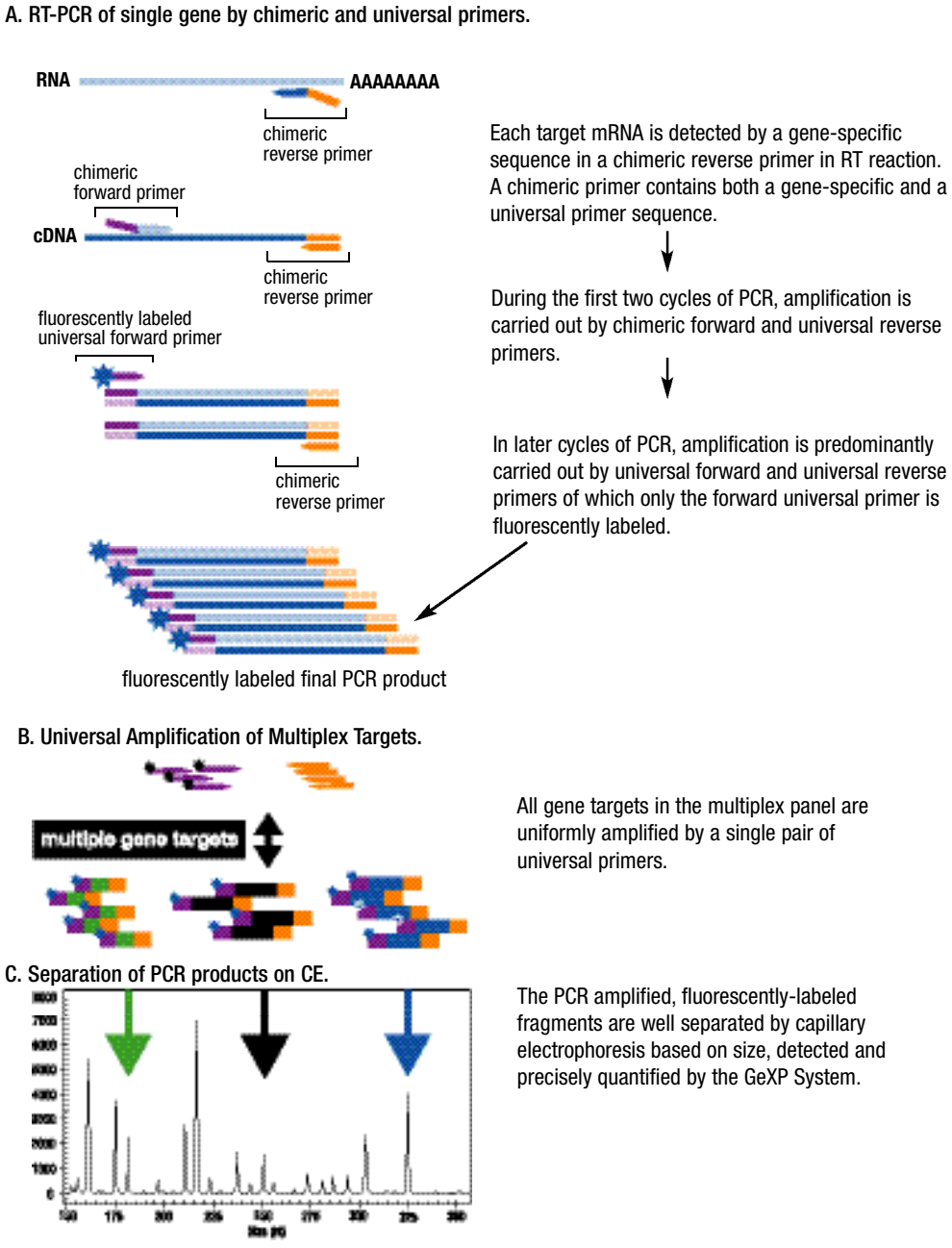

Figure 1. *GeXP Technology: multiplex universal priming strategy.*

10 min followed by 35 cycles of 94ºC 30 sec, 55ºC 30 sec, 68ºC 1 min; hold at 4ºC. To avoid possible contamination from non-specific template sources during PCR, we recommend preparing the GeXP RT and pre-PCR reactions in a location physically separated from areas where PCR and post-PCR work is being performed.

4) Separation by the GenomeLab GeXP Genetic Analysis System:

Multiplex detection via capillary electrophoresis is made possible by using chimeric primers and universal primers of which the forward universal primer is fluorescently-labeled (Figure 4). Fluorescently-labeled final PCR products are separated, detected and precisely quantified by the GeXP system (Figure 1). Following PCR, the samples are diluted with Tris-HCl buffer (Table 3), mixed with the DNA size standard-400, sample loading solution, and overlaid with mineral oil (Table 4). The separation conditions on the GeXP system are as follows: capillary temperature 50ºC, denaturation at 90ºC for 120 sec, injection for 30 sec at 2.0 kV, separation at 6.0 kV for 35 min. Note that this is the default Frag-3 protocol.

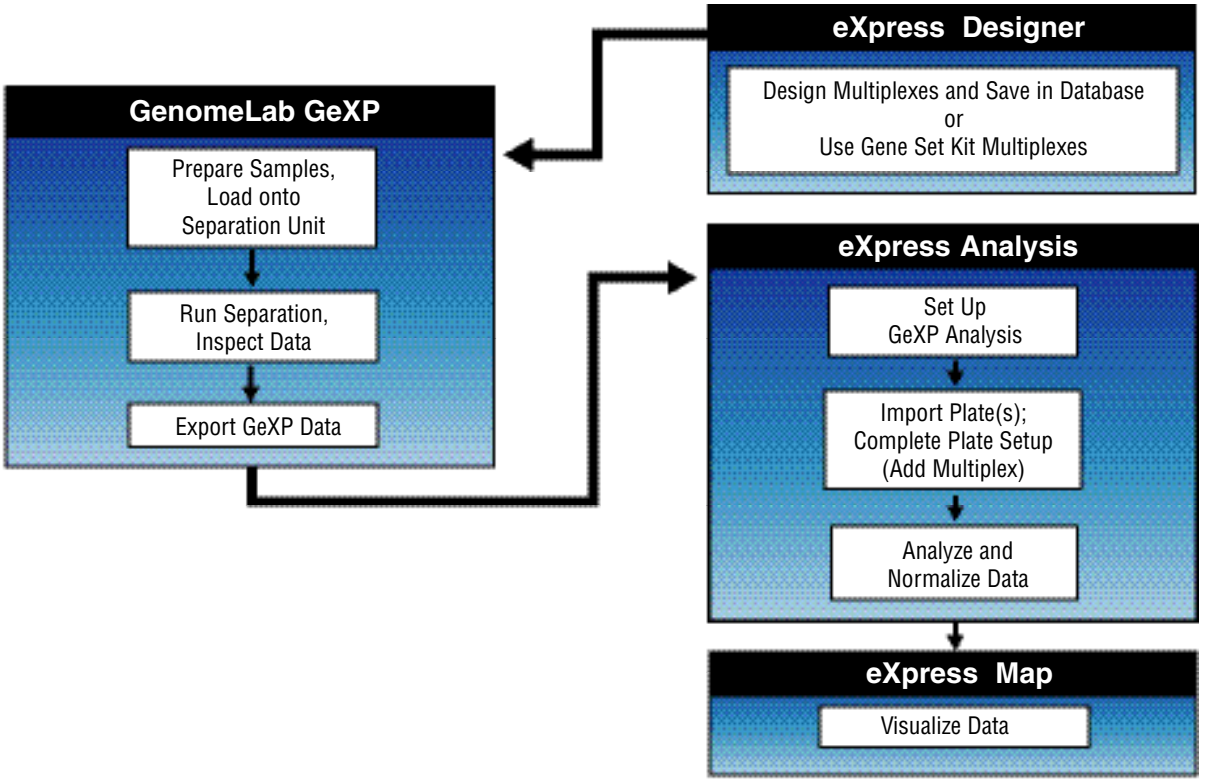

Figure 2. *Schematic of GenomeLab GeXP workflow.*

5) Fragment Analysis and Express Profiling Analysis:

Once separated, the data are initially analyzed using the Fragment Analysis module of the GenomeLab GeXP system software, followed by the eXpress Analysis module of the eXpress Profiler software (Figure 4). Raw data is first analyzed using the default GeXP analysis parameter. The fragment data, the peak height and peak area information is then imported to eXpress Analysis where i) experimental parameters such as gene set panel used, sample name, and treatment are associated with each RNA sample via Sample Layout setup, ii) experimental data and fragments are linked with gene information via Peak Binning, and iii) peak area value of each gene fragment is normalized against a reference gene and the results can be plotted and displayed. The results can be exported and subjected to 3rd party software analysis or viewed using the eXpress Map module of the

Table 2. *PCR reaction setup.*

PCR Reaction Mix	Volume per Well
PCR Buffer 5X	4.0 µL
25 mM $MgCl_2$ (ABgene)	4.0 µL
Custom PCR Fwd Primer Plex*	2.0 µL
Thermo-Start® DNA Polymerase (ABgene AB-0908/A)	0.7 µL
cDNA Samples (RT reactions from the RT Plate)	9.3 µL
Total	**20.0 µL**

***Note.** The oligo concentration for each forward primer in the PCR Fwd Primer Plex is 200 nM.*

Table 3. *Pre-dilution setup.*

Pre-Dilution Mix	Volume per Well
PCR Reaction Samples from the PCR Plate	2.0 µL
10 mM Tris-HCl pH 8.0	8.0 µL
Total	**10 µL**

**Additional 10 mM Tris-HCl pH 8.0 can be added to optimize the Sample Pre-dilution concentration.*

Table 4. *GenomeLab sample setup.*

GenomeLab Sample	Volume per Well
PCR Reaction Samples (undiluted from the PCR Plate or diluted from the Pre-dilution Plate)	1.0 µL
DNA Size Standard 400	0.5 µL
Sample Loading Solution	38.5 µL
Total	**40.0 µL**
Mineral Oil	1 drop

Table 5. *Gene list and expected size.*

Gene Number	Expected size (nt)
Ls3h2	147
DAG2	154
LEA	166
ACT2/7	173
ETR1	185
RGL2	192
COP9	200
HY1	209
APT1	221
CTS	228
ABA3	238
TUB2	246
DET3	261
Ls3h1	279
AUX1	289
KAN^r	325

Reference genes are indicated by italicized font.

eXpress Profiler software. In the eXpress Map, data can be visualized as i) Heat map table for the indication of up- or down-regulation of each gene across all samples tested, ii) Profile display of relative expression changes, iii) Correlation table for displaying pair-wise linear correlation of genes, and iv) K-means for clustering genes with similar profiles.

Validation of GeXP

An F_8 recombinant inbred line was created from a cross between an iceberg lettuce (*Lactuca sativa* L.) and a butter lettuce type. Lettuce seeds were imbibed at 20ºC under red (660 nm) or far-red (730 nm) light, conditions that promote or prevent germination, respectively. RNA was extracted from the seeds using the method described by Vicient and Delseny (1998) with slight modifications.

The quality and quantity of the RNA was determined either by absorbance readings using a spectrophotometer and gel electrophoresis or an Agilent 2100 Bioanalyzer while cDNA synthesis was performed according to Beckman GeXP instructions. Further details are given under the Experimental Details section. We designed a multiplex panel which ranged in size from 147 to 289 nucleotides and consisted of three reference genes and 12 genes associated with seed germination; hereafter this panel will be referred to as the lettuce germination panel (Table 5). Each gene was verified by analyzing the singleplex reactions with the expectation that a single product of correct size was detected (Figure 5). For each of the 15 genes in the multiplex, a single PCR product was detected in addition to the internal control peak (*KAN^r*).

We next verified that the multiplex reaction did not produce spurious fragments which may be caused by non-specificity of primers, or primer-primer interactions. A greater number of fragments than

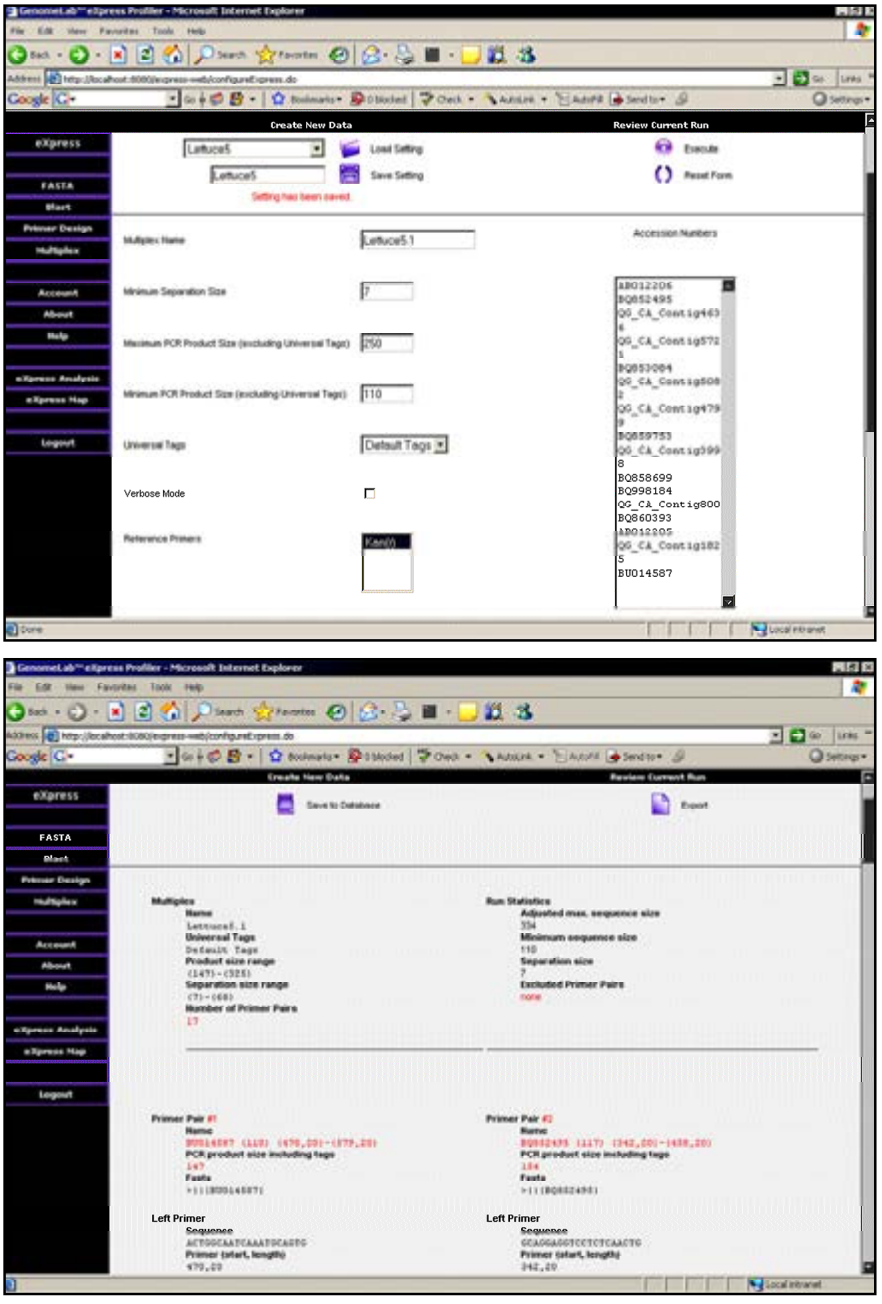

Figure 3. *Top: Input accession numbers for genes of interest. Bottom: Multiplex design output.*

expected may occur if the target gene belongs to a gene family. In circumstances where the target gene is known to be a member of a gene family, the eXpress Designer module has the flexibility to design a primer within user-defined regions of the gene. In the lettuce germination panel *LS3H2* and *DET3* were members of a multi-gene family which necessitated targeting a specific region in the gene to reduce the likelihood of targeting the paralog.

We performed GeXP on the total RNA extracted from seed imbibed under control conditions (red-light 20°C) and a total of 15 fragments were observed in the multiplex GeXP-PCR reaction (Figure 6). We used real-time qPCR as a control to determine if GeXP methodology produced essentially equivalent data.We compared the expression of genes *DAG2, LEA, HY1, COP9*, and *ABA3* using real-time qPCR with *SYBR* Green detection versus that using GeXP. After normalization to the three reference genes

(Vandesompele et al, 2003), gene *LEA* was expressed at the highest level while genes *DAG2, COP9, HY1* and *ABA3* were expressed at much lower levels in both the real-time qPCR and GeXP system (Figure 7). The normalized gene expression of *LEA* was observed to be six- and nine-fold above the expression of the reference genes using GeXP methodology and real time-qPCR, respectively. Slightly more variation was observed in the real time-qPCR analysis compared to the GeXP system. In our experience these differences between expression levels are not biologically significant and can easily arise from random variation inherent to biological systems and all quantitative methodology. We have observed that genes can usually be classified as very low, low, moderate or high expressing with each category separated by an order of magnitude. In this comparison if a given gene was placed in two different categories we would conclude the methodologies were different. This was not observed for any of the fifteen genes under any experimental condition.

Attenuation

During gene set panel development, attenuation, i.e., titration or dilution of primer concentration for

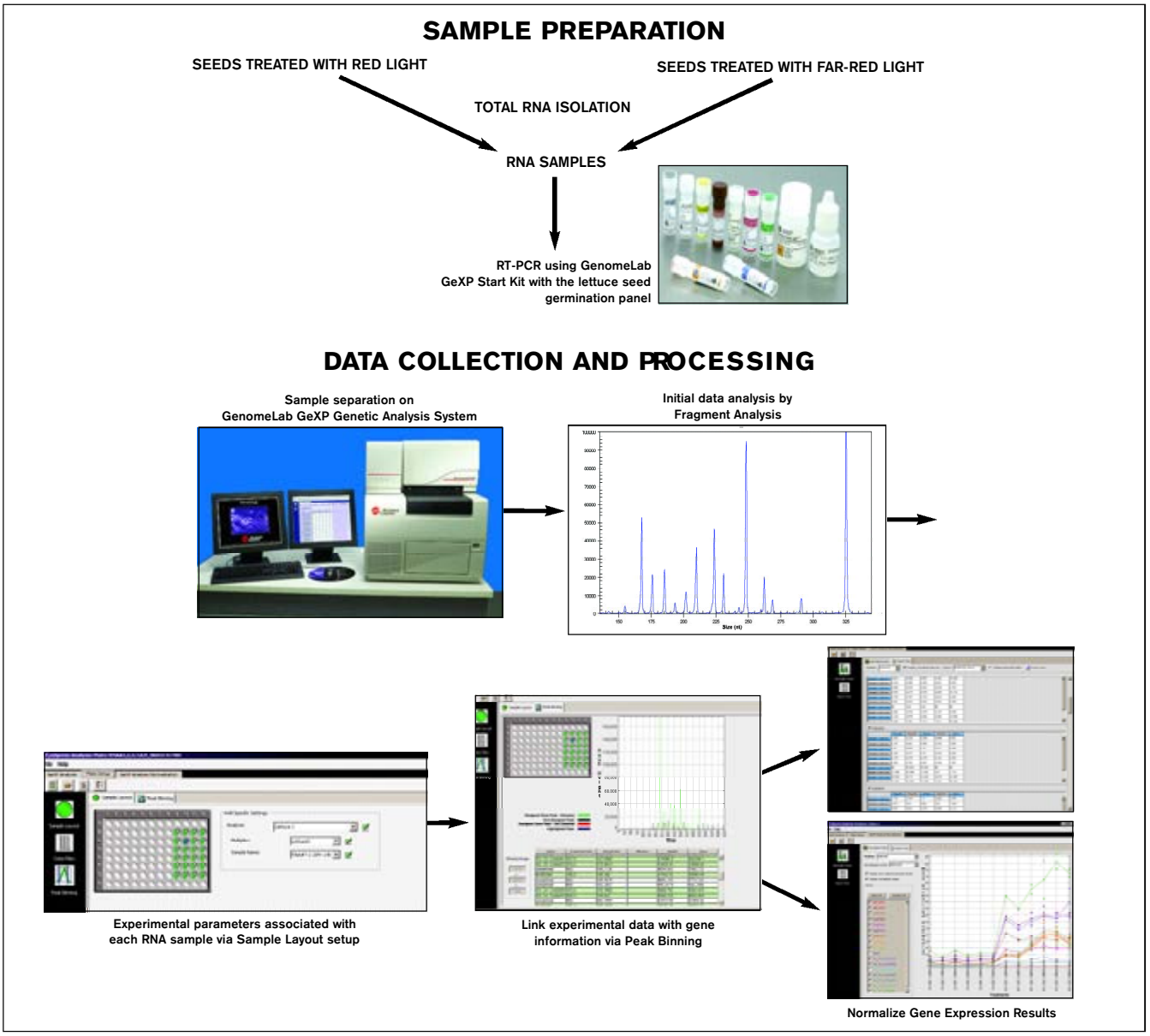

Figure 4. *Sample preparation and data processing.*

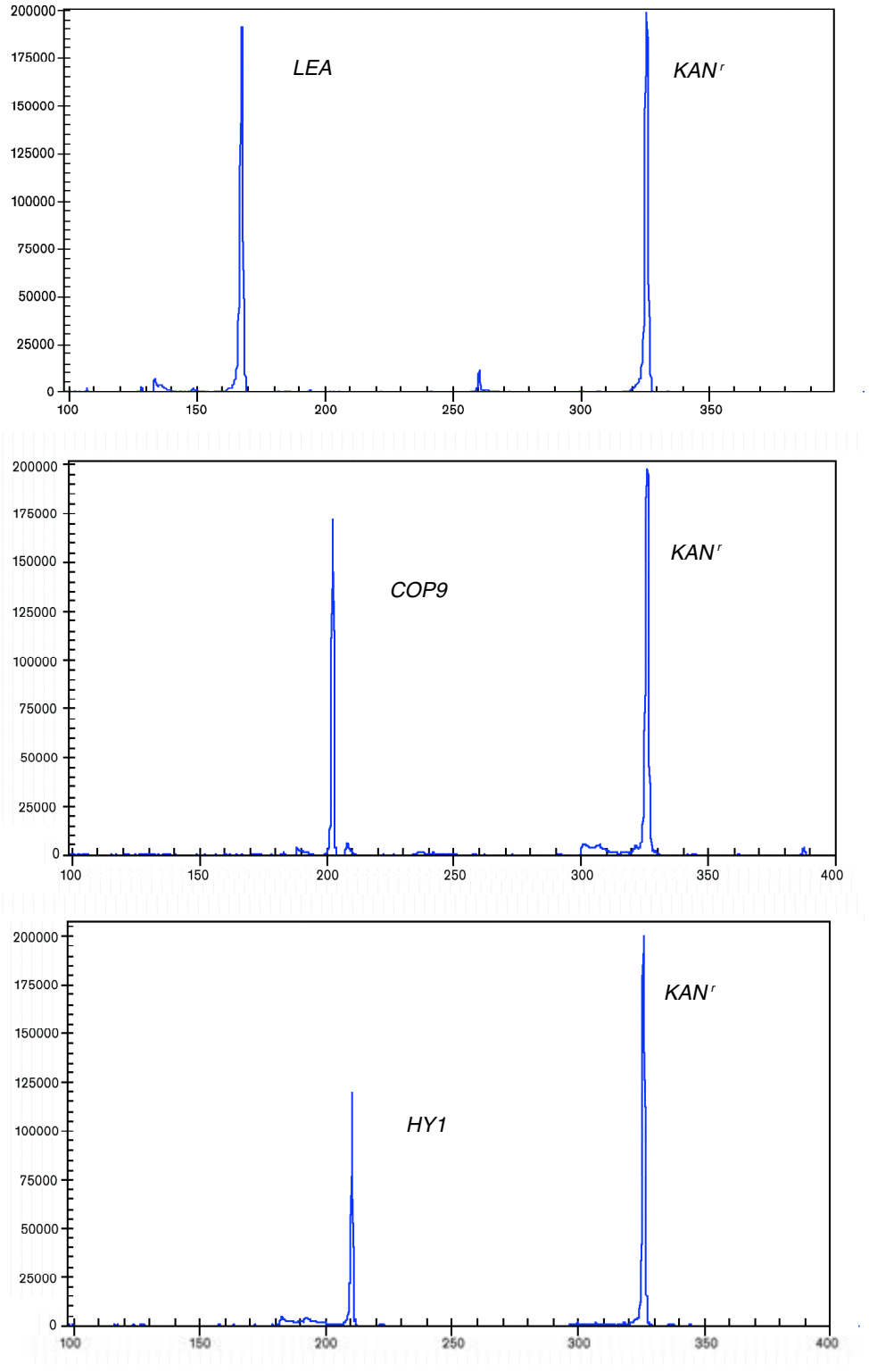

Figure 5. Singleplex primer verification for genes LEA, COP9, and HY1. *Following design of the primers the samples were processed through cDNA synthesis and PCR with a single forward primer and the full compliment of reverse multiplex primers. The targeted gene should produce a single fragment under these conditions while the KAN^r gene serves as a positive control.*

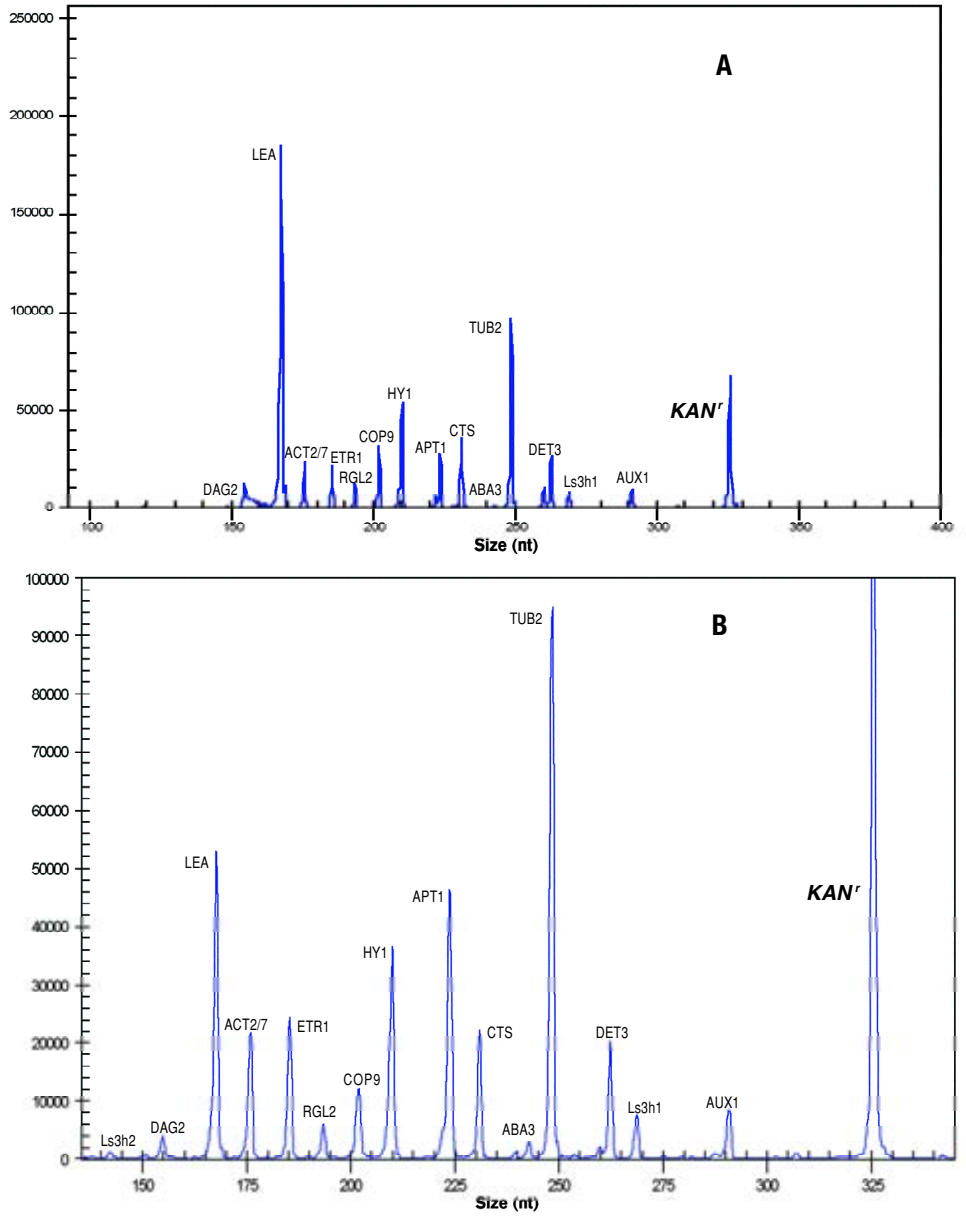

Figure 6. *Multiplex primer verification for all genes. Following design of the primers the samples are processed through cDNA synthesis and PCR with the full compliment of forward and reverse multiplex primers. A single fragment is expected for each gene under these conditions as is seen in panel (A); note that LEA is highly expressed and approaches the upper limits of detection for the GeXP system. (B). Mulitplex gene expression after attenuation of LEA.*

a certain highly expressed gene, may be necessary as part of the panel optimization. The expression level of *LEA* approached the upper limits of linearity for accurate quantification by the GeXP system and therefore required attenuation. This was accomplished by making a serial dilution series starting from 5 nM and diluting to a final concentration of 0.0390625 nM. The goal is to have the peak signal intensity of *LEA* within the mid-level of linearity for accurate quantification of the GeXP system. During capillary electrophoresis the sample is introduced into the capillary by electrokinetic injection and the number of moles introduced depends on the electroosmotic flow and electrophoretic mobility of the analyte. During injection the analytes with the highest electrophoretic mobilities will be preferentially introduced over those with lower mobilities (Jorgenson and

Lukas, 1983). Since the gene with the highest expression will be present in the cDNA pool in the highest concentration, we wanted to determine if it was preferentially injected over the others which, if true, would preclude accurate and reproducible quantification. Detection of discrimination is easily accomplished by comparing the normalized expression values of the serial dilution series.

Gene expression data were analyzed to determine if quantification of expression could be accurately calculated following dilution. A second degree polynomial equation accurately predicted the log concentration in all eight capillaries as indicated by the high coefficient of multiple determination values (range 0.87 to 0.98; Figure 8). We subjected the expression levels of the other twelve genes at each dilution series to regression analysis to detect if discrimination during injection was occurring. We would expect a flat response if discrimination was not occurring, and this was what was observed as illustrated by the expression levels of genes *COP9* and *CTS* (Figure 9). We concluded from these experiments that dilution of the attenuated gene had no effect on the ability to accurately quantify gene expression and that discrimination during injection did not occur in response to dilution.

The occasion may arise where one would like to compare the values from one capillary to another, as might occur when different biological samples are used. Although the peak areas are highest in the middle capillaries once the data are normalized against internal reference genes the expression values are comparable. Once the data are normalized the values for each gene are similar regardless of concentration of the attenuated gene, and the values for a given gene are the same between any two capillaries (Figure 10). Thus, we conclude that attenuation does not affect the quantification of gene expression and that once normalized, gene expression values are equivalent across all capillaries. The normalization validation means that comparisons can be made for a single gene across capillaries as well comparisons among different genes within the same capillary.

Gene Expression Profiling Analysis

Once the gene set is optimized and the primer concentrations were determined, the panel can be used over and over again for express profiling analysis following the GeXP protocol. The finalized lettuce germination panel was used for gene expression analysis under two different light conditions during imbibition, red or far-red light,

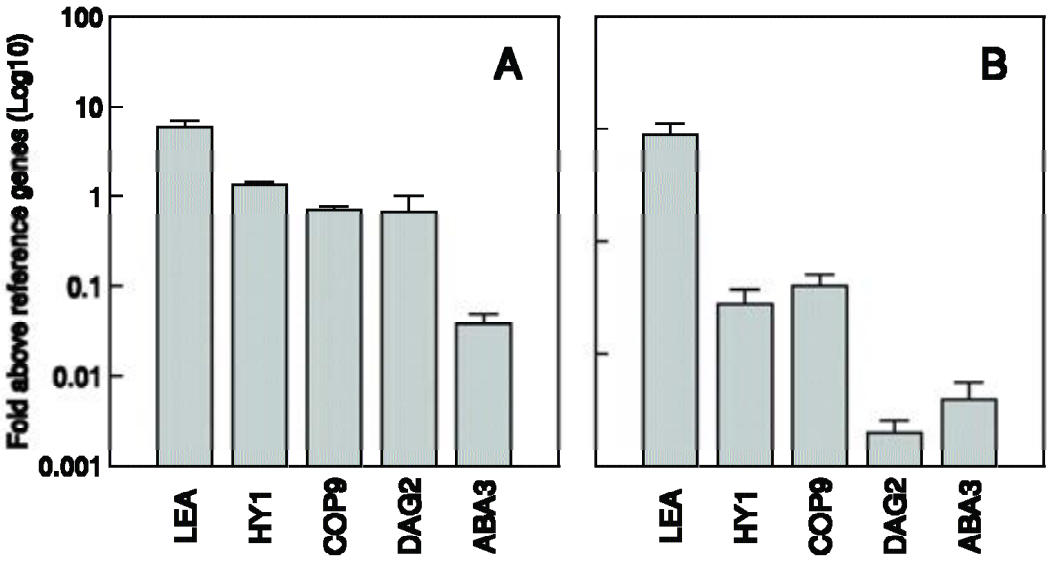

Figure 7. *A comparison of gene expression using GeXP (A) or real-time quantitative PCR (B) for a highly expressed gene LEA and four low expressing genes (DAG2, COP9, HY1, ABA3). The gene expression is normalized and expressed as the fold increase above three reference genes.*

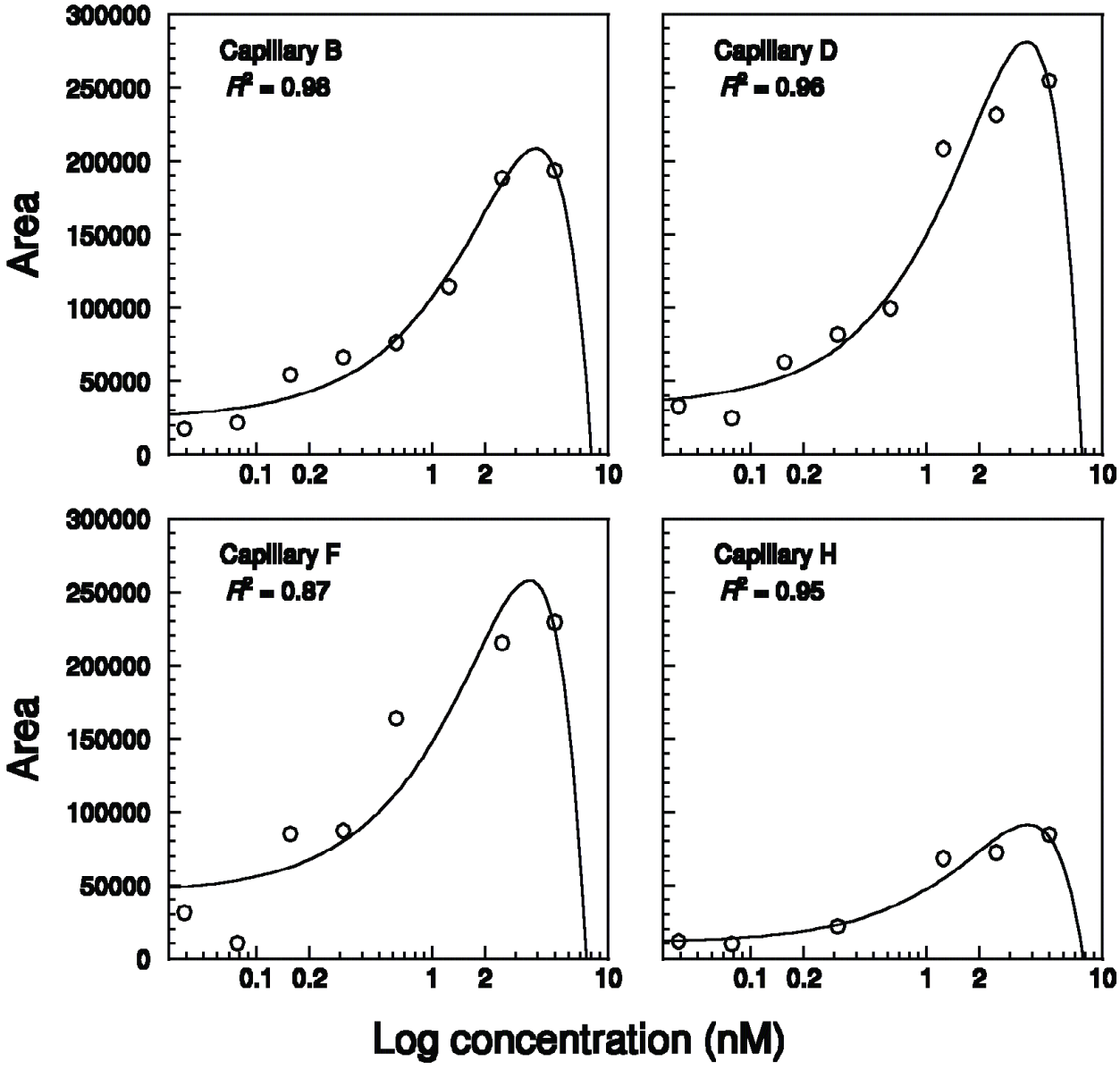

Figure 8. *Attenuation of gene LEA, a highly expressed gene. A highly expressed gene may need to be attenuated to stay within the linear range of detection. Its relative expression level can be determined from a polynomial regression equation. In the example above the primer for gene LEA was serially diluted from 5 nM to 0.0390625 nM and the sample run in capillaries B, D, F and H. The predicted area is shown by the solid line and the observed values are represented by the open circles. The coefficient of multiple determination (R^2) values were 0.98, 0.96, 0.87 and 0.95 for capillaries B, D, F and H respectively.*

that promotes or prevents germination, respectively (Figure 4). In general, the normalized GeXP expression values were slightly higher than those observed in real-time qPCR (Figure 11). There was very close agreement between the two systems in expression levels in each of genes examined, and each gene was expressed at a higher level in red light than under far-red light with the exception of gene *AUX1* which appeared to be up-regulated under far-red light. Gene *LEA* was expressed at a higher level under far-red light conditions than red-light when evaluated using real-time qPCR although little difference was observed using the GeXP system (Figure 11). Because the two

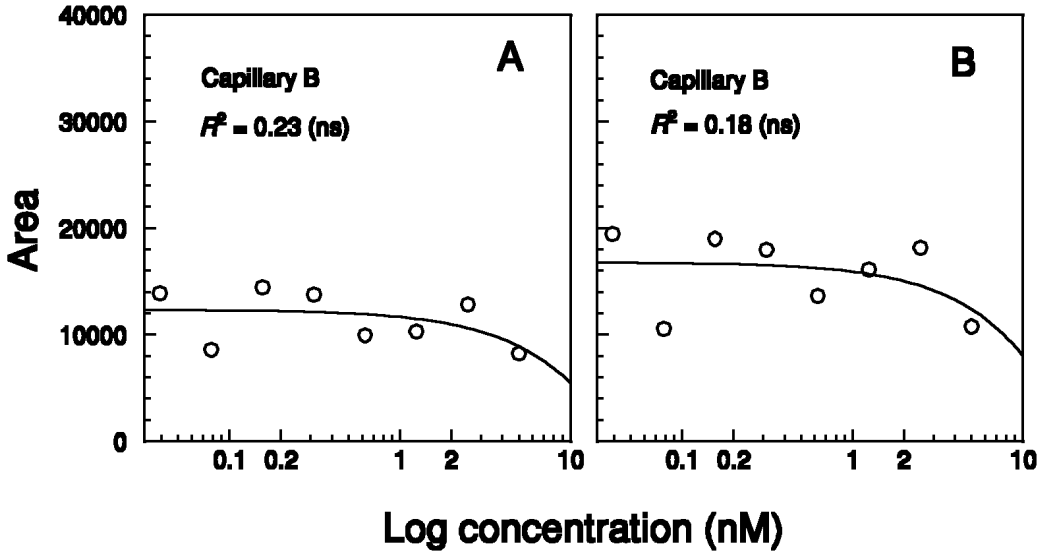

Figure 9. *Expression level of genes COP9 (A) and CTS (B) in response to attenuation of gene LEA, a highly expressed gene. Attenuation of a highly expressed gene does not have an effect on the other genes in a multiplex and quantification of gene expression is not affected.*

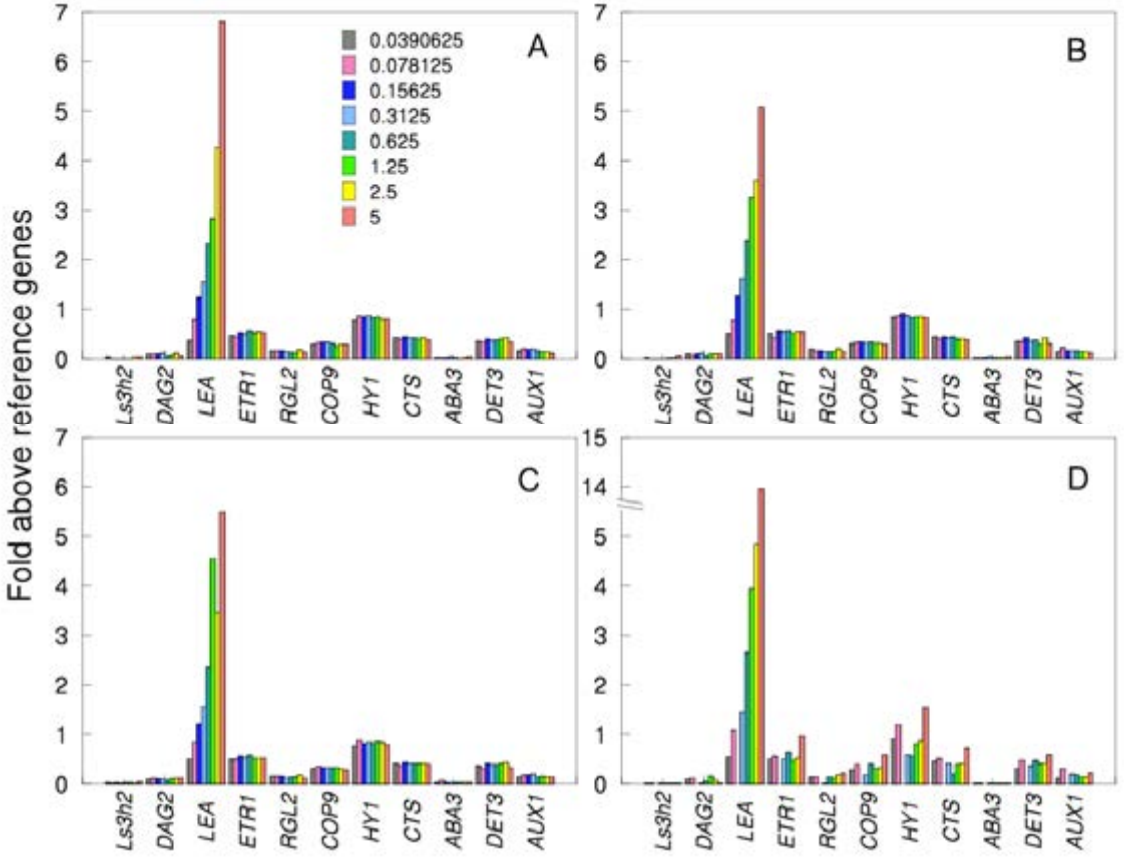

Figure 10. *Normalized expression of eleven genes in response to attenuation of a highly expressed gene (LEA). The gene expression in capillaries B, D, F, and H (A, B, C, D, respectively) after normalization to reference genes is essentially the same across all dilutions and among all capillaries. These data indicate attenuation did not affect the expression level of any of the other genes in the multiplex.*

technologies are different we would expect slight differences between the two methods but overall the data between GeXP and real-time qPCR are in close agreement.

Cost Analysis

The GeXP method is ideally suited for expression of a moderately high number of genes or a high number of samples or genotypes. The savings in time and money are realized as more genes are added to the panel. In the Still lab we have designed and validated panels with as many as 40 genes. The time required preparing samples for real time-qPCR and GeXP is approximately the same. By multiplexing the time throughput is increased by several fold. For example, the Still laboratory is working toward mapping expression level polymorphisms for 151 germination-associated genes in a recombinant inbred line containing 133 genotypes. We have calculated that with three biological replicates each with three technical replicates (i.e. wells on a 96 well plate) it would require 1,855

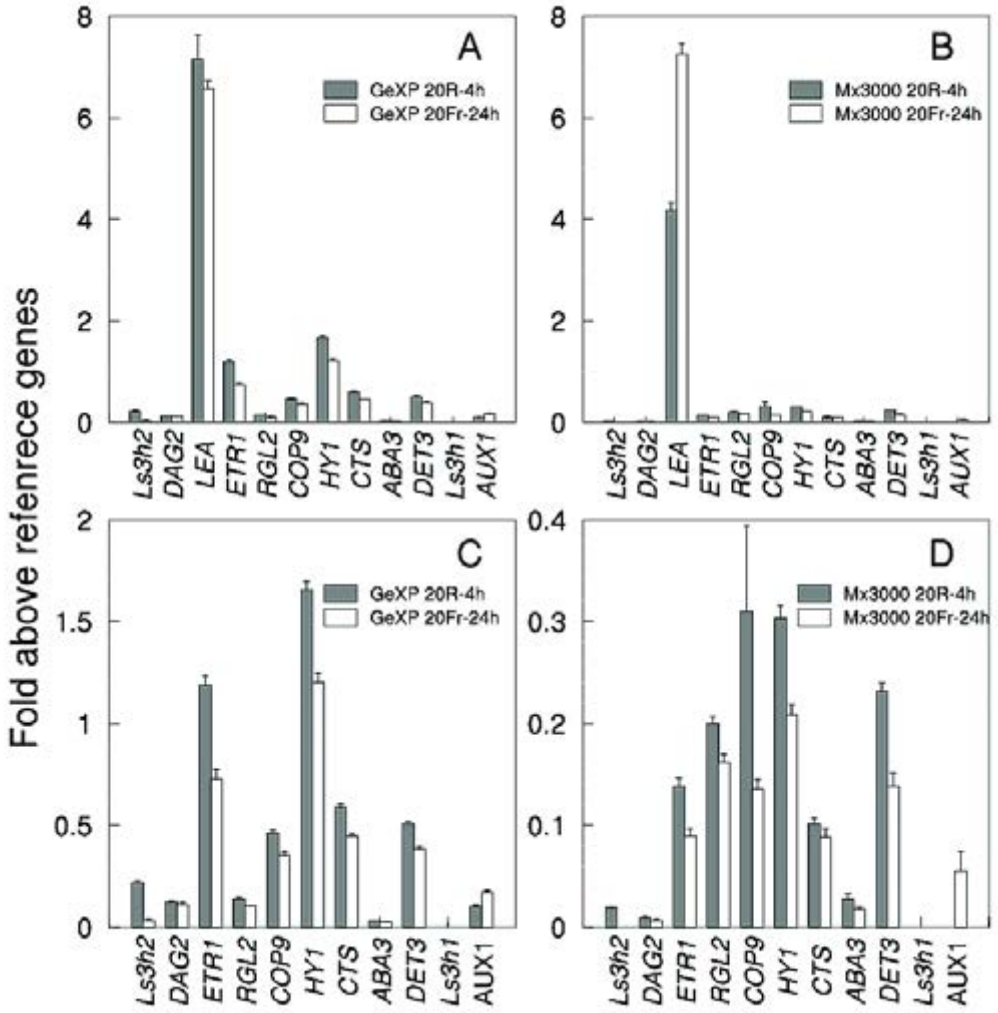

Figure 11. *A comparison of gene expression using GeXP (A, C) and real-time quantitative PCR using a Stratagene Mx3000 and SYBR Green detection (B, D). The data were normalized against the geometric average of three reference genes and represents the normalized expression relative to the three reference genes. Lettuce seeds of genotype F[8]-130 were exposed to red light (darkened bars) for four hours or far-red light (open bars) for 24 hours. Panels A and B include the results of LEA whereas the same data are presented without LEA (C, D) to facilitate viewing of low-expressed genes. Note that the data shown in these graphs do not represent attenuation of highly expressed genes.*

plates and take 618 days to complete the analysis if three plates were analyzed per day using real-time qPCR. By comparison, the same data can be generated in 11 days using the GeXP approach employing five germination panels.
We have rarely observed a sample failing during capillary electrophoresis and have observed remarkably little variation between replicates of the same sample. Therefore, technical replications are not necessary and the above analysis can be accomplished using 21 plates with a maximum of two plates analyzed on the GeXP system per day.
We next compared the reagent and consumables costs associated with real-time qPCR versus GeXP. The costs associated with real-time qPCR include the cDNA synthesis kit, real time-qPCR SYBR kit which includes Taq polymerase, and other consumables that add a small fraction to the overall cost. We have estimated a cost of one gene for one genotype (one biological replicate, three technical replications) for real-time qPCR is around $4.11. Whereas the cost to analyze a single gene using the GeXP is quite high, if 30 genes are multiplexed the cost per gene per genotype (one biological replicate) is about one-tenth that of real-time qPCR. The costs associated with GeXP include the GeXP start kit, Taq polymerase, and consumables that include an array, gel and sample loading solution, plus minor consumables.

Conclusion

In summary, we have shown that the GeXP produces quantitative gene expression data that is equivalent to real-time qPCR, and that by including reference genes in each panel the data can be normalized for each capillary. This allows one to make quantitative comparisons for a single gene within and between capillaries. The flexibility and type of data generated using GeXP is thus identical to that obtained from real-time qPCR. And finally, we calculated the time consumption of the GeXP method in analyzing a moderate number of genes across a moderate number of genotypes or samples is 20-55-fold[1] lower than real-time qPCR and the reagent and consumable costs of the GeXP method is 5-10-fold[2] lower than real-time qPCR.

*Note: [1]Fold difference calculation in time consumption depends on whether or not primer design and optimization is included in the calculation for 151 genes. [2]Fold difference calculation in cost analysis depends on how many genes are in a GeXP panel, which in this case 15 or 30 genes were used in the calculations.

References

Bigler, D, Gulding KM, Dann R, Sheabar FZ, Conaway MR, Theodorescu D (2003) Gene profiling and promoter reporter assays: Novel tools for comparing the biological effects of botanical extracts on human prostate cancer cells and understanding their mechanisms of action. Oncogene 22:1261-1272.

Casati P, Stapleton AE, Blum JE, Walbot V (2006) Genome-wide analysis of high-altitude maize and gene knockdown stocks implicates chromatin remodeling proteins in response to UV-B. Plant J 46:613-628.

Golub TR, Slonim DK, Tamayo P, Huard C, Gaasenbeck M, Mesirov JP, Coller H, Loh ML, Downing JR, Caligiuri MA, Bloomfield CD, Lander ES (1999) Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. Science 286:531-537.

Jazaeri AA, Yee CJ, Sotiriou C, Brantley KR, Boyd J, Liu ET (2002) Gene expression profiles of BRCA1-linked, BRCA2-linked, and sporadic ovarian cancers. J Natl Cancer Inst 94:990-1000.

Jorgenson JW, Lukacs KD (1983) Capillary zone electrophoresis. Science 222: 266-272.

Rozen S, Skaletsky HJ (2000) Primer3 on the WWW for general users and for biologist programmers. In: Krawetz S, Misener S (eds) Bioinformatics Methods and Protocols: Methods in Molecular Biology. Humana Press, Totowa, NJ, pp 365-386.

Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, Speleman F (2003) Accurate normalization of real-time quantitative RT-PCR by geometric averaging of multiple internal control genes. Genome Biol 3:research0034.1-0034.11.

Vicient CM, Delseny M (1998) Isolation of total RNA from Arabidopsis thaliana seeds. Analy Biochem 268:412-413.

Wang E, Miller LD, Ohnmacht GA, et al. (2002) Prospective molecular profiling of melanoma metastases suggests classifiers of immune responsiveness. Cancer Res 62:3581-3586.

Watson MA, Perry A, Budhjara V, Hicks C, Shannon WD, Rich KM (2001) Gene expression profiling with oligonucleotide microarrays distinguishes World Health Organization grade of oligodendrogliomas. Cancer Res 61:1825-1829.

Wong ML, Medrano JF (2005) Real-time PCR for mRNA quantitation. BioTechnieques 39:75-85.

All trademarks are the property of their respective owners.

www.sciex.com/ce

View SCIEX products at www.sciex.com
Find your local office at www.sciex.com/offices

AB SCIEX Headquarters
500 Old Connecticut Path | Framingham, MA 01701 USA
Phone 508-383-7700
www.absciex.com