

A high-throughput gene expression analysis technique using competitive PCR and matrix-assisted laser desorption ionization time-of-flight MS

Chunming Ding* and Charles R. Cantor*^{††}

*Bioinformatics Program and Center for Advanced Biotechnology, Boston University, 36 Cummington Street, Boston, MA 02215; and [†]Sequenom, Inc., 3595 John Hopkins Court, San Diego, CA 92121

Contributed by Charles R. Cantor, January 27, 2003

We report here an approach for gene expression analysis by combining competitive PCR and matrix-assisted laser desorption ionization time-of-flight MS. A DNA standard is designed with an artificial single nucleotide polymorphism in the gene of interest. The standard is added to the reverse transcription product before PCR. Subsequently, a base extension reaction is carried out at the single nucleotide polymorphism position, and the products are quantified by matrix-assisted laser desorption ionization time-of-flight MS. The approach is capable of relative and absolute quantification of gene expression; it is extremely sensitive (as few as five copies of DNA were quantified) and highly reproducible. It is also capable of simultaneous quantification of both alleles for heterozygotes and alternatively spliced genes. We have incorporated this technique with the homogeneous Mass Extension system (Sequenom) to create a high-throughput, automated gene expression analysis platform where a few hundred genes from 20–500 different samples can be accurately quantified per day.

DNA and RNA can be studied in two different manners. In qualitative studies, one asks which sequences are present (1, 2). In quantitative studies, one asks how much each sequence is present (3). Better, accurate, automated, high-throughput methods for RNA and DNA quantification are badly needed. Here, we present a method that is immediately usable on a commercial genotyping system (4).

Transcriptional profiling, gene expression analysis at the mRNA level, is one of the key components of functional genomics. Abnormal gene expression levels are often associated with cellular malfunctions. Thus, they are potentially disease causing. Gene expression differences between two samples have been applied to disease diagnosis and classification (5, 6). Effects of reagents or drugs on gene expression patterns have been used to test drug efficacies and to determine pharmacological mechanisms (7, 8). Correlated mRNA expression profiles under different cellular conditions have been used to predict gene functions (9, 10). DNA and RNA quantifications have also been used for detecting bacterial (11) and viral pathogens (12), as well as host–pathogen interactions (13).

Transcriptional profiling can be carried out by several methods. Traditional methods like Northern blot analyses (14) and ribonuclease protection assays (15) are generally insensitive, labor intensive, and low-throughput. DNA microarray technology, a hybridization-based method (16), has greatly facilitated large-scale gene expression study. However, DNA microarrays do not have the highest sensitivity, and a significant percentage of genes cannot be quantified. Quantitative PCR methods, most notably real-time PCR (17) and competitive PCR (18), are the highest sensitivity tools available for gene expression studies. Real-time PCR is a medium-throughput technique widely used for relative gene expression quantification. However, special care needs to be taken to optimize PCRs so that a gene of interest is amplified with similar efficiency to a housekeeping gene used as a standard. Because the gene of interest and the standard are amplified in separate reactions, care must be taken to avoid

pipetting errors and primers binding to nonspecific targets. Competitive PCR is a method where a standard and a gene of interest are coamplified in the same reaction. Because the concentration of the standard is known, the concentration of the gene can be calculated from the ratio of the resulting PCR products. This technique potentially allows absolute expression quantification. However, conventional competitive PCR suffers from a few drawbacks that have greatly limited its applications. The PCR products from the gene of interest and the standard have to differ significantly in size so that they can be readily separated by gel electrophoresis. PCR efficiencies for the two different size templates can differ by as high as 7-fold with 40 cycles of PCR (19). The commonly used detection methods such as agarose gel electrophoresis have low dynamic range, low throughput, and are ineffective for dealing with heterodimeric DNAs formed by homologous regions between a gene of interest and a standard (20), resulting in poor accuracy and reproducibility (21). Significant efforts have been spent on increasing both the accuracy and the throughput of the competitive PCR approach, with some limited success (22, 23).

Here, we report an approach for gene expression analysis by combining a unique competitive PCR design and fully automated, extremely high-throughput matrix-assisted laser desorption ionization time-of-flight (MALDI-TOF) MS detection and quantification of oligonucleotides [hereafter called real competitive PCR (rcPCR)]. In this approach (Fig. 1), a single artificial mutation is introduced to constitute a standard for a gene of interest. This mutation can be placed, when possible, at a naturally occurring single-nucleotide polymorphism (SNP) site for a gene of interest. This design will offer simultaneous quantification of the expression of both alleles of a heterozygote. Detection and quantification of both the standard and the gene of interest are carried out by the fully automated, high-throughput MassARRAY system (Sequenom) with homogeneous Mass Extension assays. This technology produces results highly consistent with those from cDNA microarrays and real-time PCR for relative gene expression analysis. It is also capable of absolute gene expression quantification with extremely high sensitivity (as few as five copies of DNA were quantitatively detected). The quantification is also PCR cycle-independent. We have seamlessly incorporated the whole approach with the MassARRAY system for detection and quantification, and the RealSNP database at www.realsnp.com for assay designs, to achieve fully automated, high-throughput gene expression analysis. Even without multiplexing, 384 genes can be quantified on a 384-format silicon chip, and up to 500 chips can be run per day on a 200k MassARRAY system. This technology has broad applications in gene expression analysis, disease diagnosis, and infectious disease agent detection and quantification.

Abbreviations: MALDI-TOF, matrix-assisted laser desorption ionization time-of-flight; rcPCR, real competitive PCR; SNP, single-nucleotide polymorphism.

^{††}To whom correspondence should be addressed. E-mail: ccantor@sequenom.com.

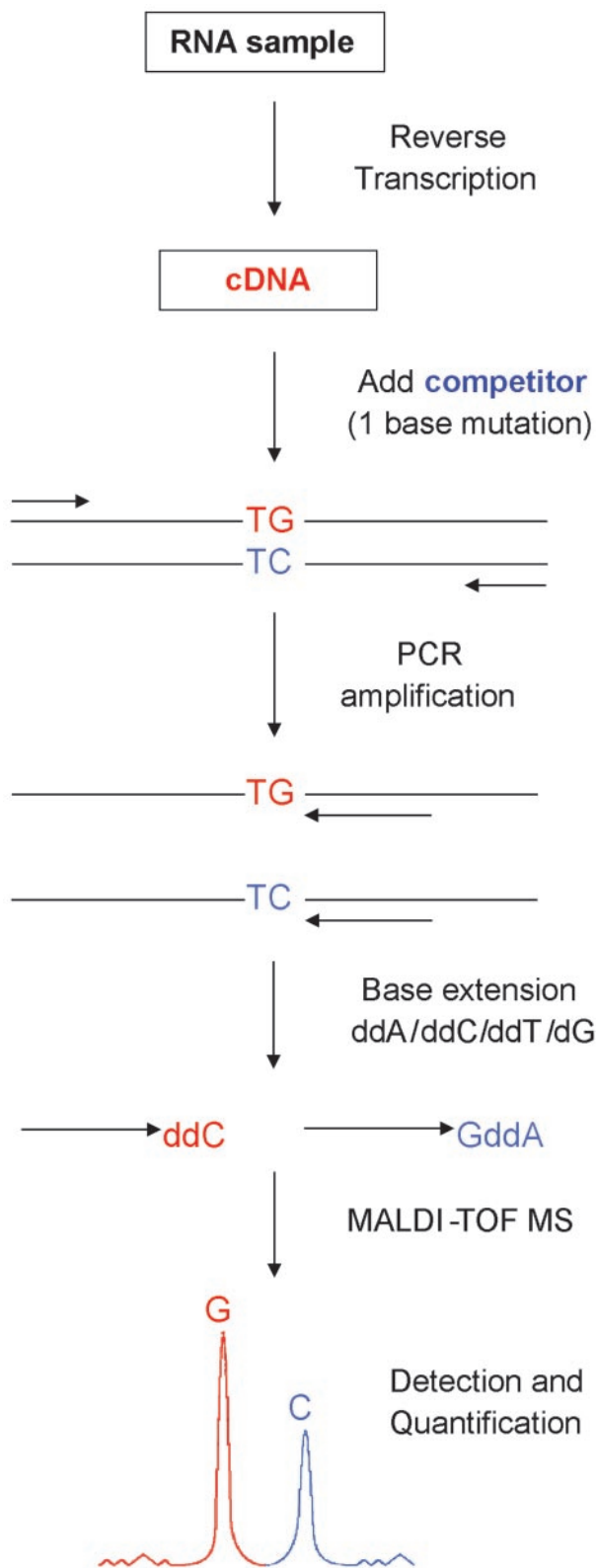


Fig. 1. The rcPCR approach for gene expression analysis. Total RNA is reverse transcribed with random hexamers. Then, a competing DNA oligonucleotide (typically 80 bases long) with one base difference from the gene of interest is added before PCR. A base extension reaction at the mutation site is carried out by adding a ThermoSequenase, three ddNTPs, and one dNTP to produce two oligonucleotide products with different molecular weights. These two products are subsequently detected and quantified by MALDI-TOF MS with ALLELO-TYPING software (Sequenom).

Materials and Methods

DNA Constructs and Primer Designs. The plasmid DNAs (wild type and mutant) for the DNA mixture experiment were kindly provided by Daniel Oprian (Brandeis University, Waltham, MA). The mutant carries an AAG→GAG mutation [Lys-296→Glu (K296E) mutation for the inserted bovine rhodopsin gene], and this mutation was used as the standard for competitive PCR. For all of the experiments, the PCR primers have a 10-base tag (5'-ACGTTGGATG-3') so that they will not interfere in mass spectra.

All DNA oligonucleotides were purchased from Integrated DNA Technologies (Coralville, IA). For human gene expression studies, the following five genes were chosen: 18s rRNA (GenBank accession no. X03205), gene A (accession no. L08246), gene B (accession no. M54894), gene C (accession no. M20681), and gene D (accession no. M59465).

PCR primers (10-base tag excluded) were as follows: for the plasmid DNAs, 5'-TTTGGGCCATCTTCATGAC-3' and 5'-TCATGATGTAGATGACCGGG-3'; for 18s rRNA, 5'-GC-GAAAGCATTGCGCAAG-3' and 5'-ACGACGGTATCT-GATCGTC-3'; for gene A, 5'-GTCACTACCCTCGACGC-3' and 5'-GAGAGATAATCTCCAGCGAC-3'; for gene B, 5'-GCAGGACATGACAACCTCATC-3' and 5'-CCATGCTA-CATTTGCCGAAG-3'; for gene C, 5'-AGTGCAAGGATGT-CACAAG-3' and 5'-AAATGATGATGGGCTGTGCG-3'; and for gene D, 5'-TCAAGGAAACAGACACACGC-3' and 5'-GCC-CCGTTTCAACAAATTCC-3'. Extension primers were as follows: for plasmid DNAs, 5'-CATCCCGGCTTCTTTGCC-3'; for 18s rRNA, 5'-TCAAGAACGAAAGTCGG-3'; for gene A, 5'-AGCGACTGCCGGTACAA-3'; for gene B, 5'-CGCAGCTTTA-AGGAGTT-3'; for gene C, 5'-GCTGGACACTCTAAA-GAGCTC-3'; and for gene D, 5'-CGCTGGCAACTGGAGTC-3'. DNA standards (typically ≈80 bases long) used as the competing templates for competitive PCR were as follows: for 18s rRNA, 5'-CTACGACGGTATCTGATCGTCTTCGAACCGCCGAC-TTTCGTTCTTGATTAATGAAAACATTCTTGGCAAAT-GCTTTTCGCT-3'; for gene A, 5'-CCGAGAGATAATCTCCA-GCGACTGCCGGTACAACGCGTCCTCCTCCTCCTGCTGGCGGGCGGCTCGAGGGTAGTGACC-3'; for gene B, 5'-GCCCATGCTACATTTGCCGAAGAGCCCTCAGGCTGG-ACTGCATAAAGCTCCTTAAAGCTGCGCAGAATGAGAT-GAGTTGTCATGTCCTGCAG-3'; for gene C, 5'-AATGATG-ATGGGCTGTCGGTAGCTGGACACTCTAAAGAGCTCT-CGCACGGTGACTTGCTTTTCTTGTGACATCCTTGAC-3'; and for gene D, 5'-CCCCGTTTCAACAAATTTCCTGAGATT-TGATAGACTCCAGTTGCCAGCGGAATTTAAAGTTGC-GTGTGTCTGTTTCCTTGA-3'. These DNA standards were PAGE-purified by the provider (Integrated DNA Technologies) and are the reverse complement of the RNA sequence, because in the first round of competitive PCR, they will compete with the cDNAs.

Reverse Transcription. A total of 100 ng of total RNA (prepared by J. Tullai, M. Schaffer, and G. Cooper, Boston University) was reverse transcribed by an AMV reverse transcriptase (ImProm-II, Promega) with random hexamers (0.25 μg in a 10-μl total reaction volume) at 42°C for 1 h. Alternatively, gene-specific primers (0.5 μM final concentration in a 10-μl total reaction) were used with a thermostable AMV reverse transcriptase (ThermoScript, Invitrogen), and reverse transcriptions were carried out for 1 h at 65°C.

PCR Amplification. HotStar *Taq* Polymerase (Qiagen, Valencia, CA) was used for all PCRs. PCR primers were at 200 nM final concentrations for a PCR volume of 5 μl. For plasmid DNA mixture experiments, a total of 10⁻⁶ μg of DNA was added. For total RNA experiments, reverse-transcribed cDNA from 5 ng of

total RNAs was added for all genes except 18s rRNA, where 0.2 ng was added. The PCR condition was: 95°C for 15 min for hot start, followed by denaturing at 94°C for 20 sec, annealing at 56°C for 30 sec, extension at 72°C for 1 min for 45 cycles (except for the plasmid DNA mixture experiment where 20, 30, or 40 cycles were performed), and, finally, incubation at 72°C for 3 min. All of these are standard homogeneous Mass Extension assay protocols (Sequenom).

Base Extension. PCR products were treated with shrimp alkaline phosphatase (Sequenom) for 20 min at 37°C first to remove excess dNTPs. A ThermoSequenase (Sequenom) was used for the base extension reactions. In contrast to standard homogeneous Mass Extension protocol, extension primers were 1.2 μM final concentration in 9-μl reactions. The base extension condition was: 94°C for 2 min, followed by 94°C for 5 sec, 52°C for 5 sec, and 72°C for 5 sec for 40 cycles. All reactions (reverse transcription, PCR amplification, and base extension) were carried out in a GeneAmp 9700 thermocycler (Applied Biosystems).

Liquid Dispensing and MALDI-TOF MS. The final base extension products were treated with SpectroCLEAN (Sequenom) resin to remove salts in the reaction buffer. This step was carried out with a Multitek (Beckman Coulter) 96-channel autopipette and 16 μl of resin/water solution was added into each base extension reaction, making the total volume 25 μl. After a quick centrifugation (2,000 rpm, 3 min) in an Eppendorf Centrifuge 5810, ≈10 nl of reaction solution was dispensed onto a 384-format SpectroCHIP (Sequenom) prespotted with a matrix of 3-hydroxypicolinic acid (3-HPA) by using a SpectroPoint (Sequenom) nanodispenser. A modified Biflex MALDI-TOF mass spectrometer (Bruker, Billerica, MA) was used for data acquisitions from the SpectroCHIP. Each matrix pad on the SpectroCHIP corresponds to one gene expression assay (except for the triplex assay, where three gene expression assays were on one matrix pad) so a 384-format chip can be used for analyzing 384 gene expressions.

MALDI-TOF MS Data Analysis. Mass spectrometric data were automatically imported into the SpectroTYPER (Sequenom) database for automatic analysis, i.e., noise normalization and peak area analysis. Systematic errors from base extension reactions caused by product inhibition and difference in desorption/ionization efficiency for the two base extension products (although these two only differ by one or two bases at the 3' end) from the DNA standard and the gene of interest were subsequently corrected with inhouse software. The correction algorithms were based on computational simulations and >25,000 individual allelotyping results carried out by Sequenom on genomic SNPs (unpublished data).

Results

Quantitative Analysis of Known DNA Mixtures. Purified plasmid DNA constructs were used to test the ability of rcPCR for DNA quantification. Two plasmid DNAs identical in sequence except one base were mixed at different ratios (10:1, 3:1, 1:1, 1:3, and 1:10) with a constant total concentration of 2×10^{-7} μg/μl. The DNA mixtures were amplified by PCR with different cycle numbers (20, 30, and 40). Shrimp alkaline phosphatase (SAP) treatment, base extension, and MALDI-TOF MS were subsequently carried out. Short oligonucleotide products from base extension reactions were readily separated and quantified by MALDI-TOF MS and the ALLELOTYPING software package. The results (Fig. 2) show excellent linearity over a 100-fold range of DNA ratios, and the measured DNA concentration ratios are consistent with DNA concentration ratios calculated by absorbance at 260 nm. Notably, the data are clearly PCR cycle number-independent. This result is expected for two virtually

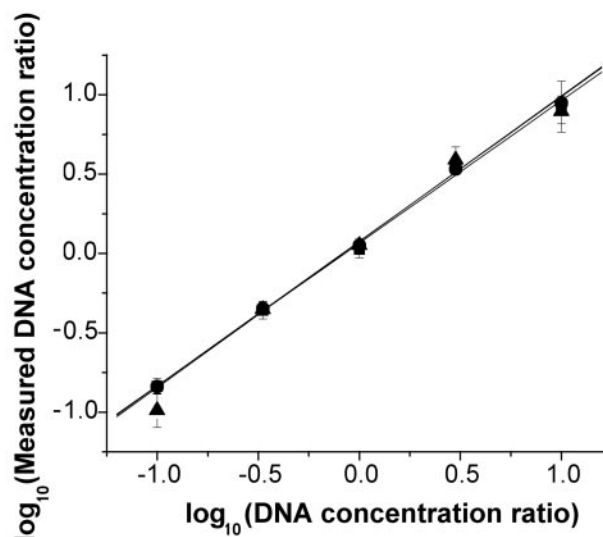


Fig. 2. PCR cycle-independent quantification of DNA concentration ratios by rcPCR. Two DNA plasmids were mixed at different ratios (1:1, 3:1, 10:1, 1:3, 1:10) but at a fixed total concentration (2×10^{-7} μg/μl). The DNAs were amplified by PCR for 20 (▲), 30 (■), and 40 (●) cycles, respectively, base extension (40 cycles) was performed, and the DNAs were spotted onto a silicon chip pre-spotted with a matrix of 3-hydroxypicolinic acid (3-HPA) and analyzed with MALDI-TOF MS.

identical templates, because two “identical” templates would exhibit the same amplification kinetics. This PCR cycle number independence makes it unnecessary to stop the PCR at the exponential stage, which is an absolute requirement for most quantitative PCR techniques (real-time PCR excluded) and greatly simplifies the procedure for gene expression analysis. In subsequent experiments, 45 cycles of PCR were used as the standard procedure.

Relative Gene Expression Analysis. Next, we tested this technique on gene expression from cultured human cells. DNA standards (≈80 bases long) were synthetic oligonucleotides with sequence identical to a portion of the cDNA sequence of the genes of interest, except with a single base mutation roughly in the middle of the oligonucleotides. The DNA standards were purified by PAGE and quantified by absorbance at 260 nm. The DNA standards then were added to the reverse-transcribed total RNA samples at 100-fold series dilutions (three dilutions, covering a 10^6 dynamic range). Competitive PCR and the MALDI-TOF MS were used for gene expression analysis. For relative gene expression analysis, the gene expression values of cells stimulated with growth factors under several conditions (preparations 1–4) were normalized to the untreated cells and are summarized in Table 1. Three reverse transcription reactions with random hexamers were carried out on the same RNA samples, and SDs were calculated. Typical coefficients of variation (CV) are <10%

Table 1. Relative gene expression analysis by rcPCR

Treatment	Gene A	Gene B	Gene C	Gene D
1	1.48 ± 0.02	10.0 ± 0.98	3.24 ± 0.82	5.82 ± 1.57
2	1.67 ± 0.23	0.68 ± 0.05	6.88 ± 0.38	0.85 ± 0.34
3	1.02 ± 0.10	6.95 ± 0.46	1.06 ± 0.06	5.08 ± 0.88
4	0.87 ± 0.06	0.29 ± 0.02	1.29 ± 0.14	0.81 ± 0.05

Expression values in the table are for cells treated with different reagents normalized to untreated cells; mean ± SD for three independent reverse transcription reactions.

Table 2. Absolute gene expression analysis with rcPCR

Absolute gene expression level	18s rRNA, $\times 10^{-1}$	Gene A, $\times 10^{-6}$	Gene B, $\times 10^{-7}$	Gene C, $\times 10^{-6}$	Gene D, $\times 10^{-7}$
Treatment* μM in 1 $\mu\text{g}/\mu\text{L}$ total RNA					
Untreated	2.27 \pm 0.19	2.07 \pm 0.14	2.99 \pm 0.05	1.09 \pm 0.11	0.79 \pm 0.18
1	2.57 \pm 0.21	3.46 \pm 0.32	33.9 \pm 2.97	3.98 \pm 1.12	5.21 \pm 1.96
2	2.38 \pm 0.22	3.61 \pm 0.47	2.14 \pm 0.33	7.81 \pm 0.19	0.66 \pm 0.14
3	2.18 \pm 0.25	2.01 \pm 0.03	19.9 \pm 0.48	1.10 \pm 0.06	3.75 \pm 0.48
4	2.42 \pm 0.12	1.94 \pm 0.22	0.93 \pm 0.08	1.49 \pm 0.10	0.67 \pm 0.10
Treatment gene copies per 1 million copies of 18s rRNA					
Untreated	10^6	9.12 \pm 0.26	1.32 \pm 0.12	4.79 \pm 0.38	0.34 \pm 0.05
1	10^6	13.5 \pm 0.55	13.2 \pm 0.18	15.4 \pm 3.3	2.01 \pm 0.65
2	10^6	15.2 \pm 2.48	0.90 \pm 0.12	33.0 \pm 3.7	0.28 \pm 0.08
3	10^6	9.32 \pm 0.94	9.19 \pm 0.86	5.05 \pm 0.31	1.72 \pm 0.12
4	10^6	7.98 \pm 0.59	0.39 \pm 0.05	6.13 \pm 0.20	0.28 \pm 0.03

The absolute numbers are represented by either micromolar gene concentration in 1 $\mu\text{g}/\mu\text{L}$ total RNA or gene copy numbers relative to one million copies of 18s rRNA; mean \pm SD for three independent reverse transcription reactions.

*Treatment with different reagents as listed.

from three independent reverse transcription reactions (Table 1). Typical CVs for four competitive PCRs from the same reverse transcription product are $<3\%$ (data not shown). The relative expression data are comparable with those from microarray and real-time PCR analysis on the same RNA samples (J. Tullai, M. Schaffer, and G. Cooper, unpublished data). For gene B and 18s rRNA, we also used gene-specific primers and a thermostable reverse transcriptase for doing reverse transcription at 65°C , and the results were consistent with those with random hexamers (data not shown).

Absolute Gene Expression Analysis. We also explored the ability of this technique to provide absolute gene expression analysis. Because we know the absolute amount of DNA standard added into the sample, and because the amplification and detection process preserves faithfully the ratio of the DNA standard and the gene of interest, we reasoned that we also should be able to obtain the absolute quantity of the transcript of interest. Absolute gene expression data are presented (Table 2) in two ways. First, gene expression levels are presented as μM concentrations in 1 $\mu\text{g}/\mu\text{L}$ total RNA. Ribosomal RNAs are the predominant RNA in total RNA samples, typically equal to $\approx 80\%$; 18s rRNA represents about 1/3 of all rRNAs. Thus, roughly about 1/4 of total RNA is 18s rRNA. By using this assumption, the concentration of 18s rRNA in 1 $\mu\text{g}/\mu\text{L}$ total RNA is $\approx 0.41 \mu\text{M}$. Thus, we estimated that the reverse transcription efficiency for 18s rRNA is $\approx 58\%$. Our analysis of 18s rRNA at five different conditions (Table 2) showed that 18s rRNA is expressed at almost constant levels ($0.238 \pm 0.012 \mu\text{M}$ in 1 $\mu\text{g}/\mu\text{L}$ total RNA), suggesting that 18s rRNA is a good candidate for normalization. This allowed us to compute the absolute gene expression as the copy number of a specific gene relative to one million copies of 18s rRNA (Table 2). This is calculated by dividing the concentration of the gene of interest by the concentration of 18s rRNA.

Sensitivity of rcPCR. A highly sensitive gene expression analysis technology is advantageous for measuring genes expressed at low levels or when sample source is limited. Gene D was selected for testing the sensitivity of the rcPCR approach. The transcript copy number of gene D was first quantified, and the same amount of its DNA standard was added. The mixtures then were diluted to contain 5, 10, or 50 copies of the DNA standard and gene D for rcPCR analysis. The result (Fig. 3) shows that our approach can quantitatively detect gene expression for as few as five copies in a 5- μL reaction sample before PCR. The internal DNA standard can also be used to quality-control the PCR amplification reactions, because an inefficient PCR will not

produce a signal in the mass spectrum. Because a successful detection of gene expression will rely on both PCR primers and the extension primer efficiently binding to the template to produce extension products of the correct molecular weights, our approach is more specific than traditional competitive PCR methods.

Assay Multiplexing. To increase further the throughput and reduce cost for large-scale gene expression analysis, we tested gene expression quantification with a triplex PCR. Genes A, B, and C were coamplified with their respective standards in the same reaction for PCR and primer extension. The extension products (Fig. 4) were clearly separated in the mass spectrum and quantified by their peak areas. It is evident from the mass spectrum that high levels of multiplexing can be achieved because triplexing led to no significant deterioration in the precision of the method.

Discussion

Gene expression analysis is of fundamental importance in understanding cellular functions. Most current techniques have focused mainly on relative transcription levels. We have combined competitive PCR and MALDI-TOF MS and incorporated this into the

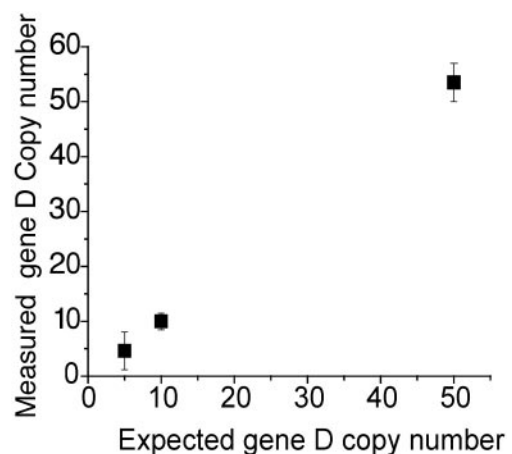


Fig. 3. Sensitivity of the rcPCR approach. Gene D was analyzed to determine its cDNA concentration and then diluted so that 5, 10, and 50 copies were added to the PCRs. DNA standard for the gene was added into the PCRs at 5, 10, and 50 copies, correspondingly. Four separate PCRs were carried out for each data point; SDs are shown.

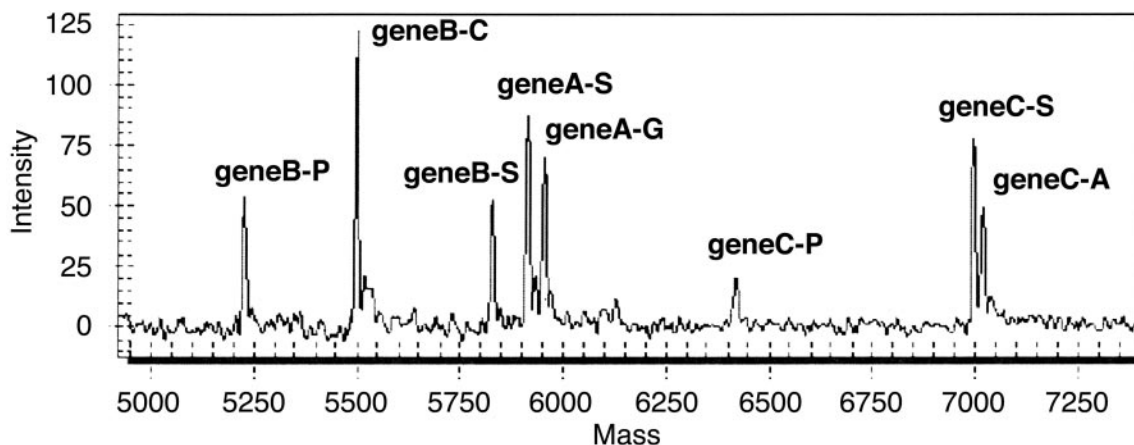


Fig. 4. Multiplexed rcPCR. Genes A, B, and C and their respective standards were coamplified in the same reaction; the mass spectrum is shown. P, the unextended extension primer; S, the extended oligonucleotide from the DNA standard; and C, G, and A, the specific natural alleles of the three genes. The quantitative results by multiplex reactions agree well with those from uniplex reactions.

MassARRAY system to create a high-throughput, fully automated gene expression analysis technology, rcPCR. The use of a DNA standard with a single point mutation enables virtually same PCR-amplification efficiency for the gene of interest and the standard. Traditional competitive PCR uses standards of different sizes compared with their genes of interest, and the same PCR efficiency cannot always be achieved. A point mutation that either creates or eliminates a restriction enzyme site was proposed for designing DNA standards (18). However, heterodimeric DNAs resistant to enzyme digestion complicate this procedure, making it unsuitable for high-throughput analysis. Our detection technique with base extension and MALDI-TOF MS completely eliminates the heterodimeric DNA problem. The two products from the gene of interest and the DNA standard are unequivocally detected and quantified. Because the gene of interest and the DNA standard are coamplified in the same PCR, the method is much less sensitive to artifacts like primer mispriming, unspecific amplification, and pipetting errors. Our measurements on the same reverse-transcribed product have CV typically <3% for direct transcriptional profiling with uniplex assays. Existing array methods can measure a large number of genes; however, the data are generally noisy and only large expression changes (e.g., >2-fold) can be confidently detected. Because of the extremely high accuracy and reproducibility, we will be able to detect small, yet biologically significant expression changes.

If a naturally occurring SNP site is used for designing DNA standards, one can simultaneously quantify the expression of both alleles of heterozygotic individuals. The DNA standard can be designed to have a mutation that carries the third (or fourth) possible base at the natural SNP position. For example, if the natural SNP is an A/C polymorphism, a T or G base at the SNP position can be used in the DNA standard. Although all SNP genotyping efforts currently have been focusing on genomic DNA level, our approach enables us to genotype at the mRNA level. This ability may represent a substantial advantage for disease association studies, because mutation and expression data can be combined (24). We can also study expression of alternatively spliced mRNAs with this technique. One way of doing that is to design quantification assays targeting each of the individual exons of a gene and compare the absolute expression levels of different exons. A significantly lower expression level of an exon is indicative of this exon being discarded in RNA splicing.

The competitive PCR approach makes absolute gene expression analysis possible because the exact copy numbers of the DNA standards added to the PCRs can be determined accurately. Currently, serial analysis of gene expression (SAGE; ref. 25) is commonly used for absolute quantification of gene expression. However, SAGE is significantly limited in quantifying rare transcripts. Our technique is extremely sensitive and will be advantageous in rare transcript quantification. Normalization with a housekeeping gene is crucial for such analyses. In our case, we found 18s rRNA equally expressed under five different treatments (CV \approx 5%). Thus, it is an ideal standard for normalization. The best way to represent absolute gene expression is relative to a housekeeping gene RNA copy number, because this will normalize for the RNA preparation and, to some extent, the reverse transcription efficiency. Absolute DNA/RNA quantification is advantageous in many aspects. First, data obtained from various sources can be compared directly. This comparison can be used to standardize and centralize gene expression analysis data. Second, the same approach can be used to quantify infectious agents like virus and bacteria, exploiting the extremely high sensitivity of our technique. The high precision of rcPCR suggests that this method will also be useful for the absolute DNA quantification necessary for diagnosis of trisomies, monosomies, loss of heterozygosity, and gene amplification.

We have seamlessly incorporated the rcPCR technique into the MassARRAY system for highly automated gene expression analysis. For a typical 384-format SpectroCHIP (see *Materials and Methods* for details), we can analyze 384 genes with uniplex assays (or \approx 2,000 genes for five-plex assays). Between 20 and 500 chips can be analyzed per day per MassARRAY system. This technique will allow academic and industrial researchers the opportunity for accurate, high-throughput gene expression studies. Our technology is ideal in these settings where it is required to analyze hundreds of different patient samples on the order of a few hundred genes.

We thank Dr. John Tullai, Mr. Michael Schaffer, and Dr. Geoffrey Cooper (Boston University) for providing total RNA samples and microarray and real-time PCR data for comparison; and Ms. Shengnan Jin and Dr. Daniel Oprian (Brandeis University) for providing the DNA constructs for the DNA mixture experiment. This work was supported by a grant from Sequenom to Boston University.

1. Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., et al. (2001) *Nature* **409**, 860–921.

2. Venter, J. C., Adams, M. D., Myers, E. W., Li, P. W., Mural, R. J., Sutton, G. G., Smith, H. O., Yandell, M., Evans, C. A., Holt, R. A., et al. (2001) *Science* **291**, 1304–1351.

3. Roth, C. M. (2002) *Curr. Issues Mol. Biol.* **4**, 93–100.
4. Jurinke, C., van den Boom, D., Cantor, C. R. & Koster, H. (2001) *Methods Mol. Biol.* **170**, 103–116.
5. Bittner, M., Meltzer, P., Chen, Y., Jiang, Y., Seftor, E., Hendrix, M., Radmacher, M., Simon, R., Yakhini, Z., Ben-Dor, A., *et al.* (2000) *Nature* **406**, 536–540.
6. Golub, T. R., Slonim, D. K., Tamayo, P., Huard, C., Gaasenbeek, M., Mesirov, J. P., Coller, H., Loh, M. L., Downing, J. R., Caligiuri, M. A., *et al.* (1999) *Science* **286**, 531–537.
7. Namba, H., Iwadata, Y., Kawamura, K., Sakiyama, S. & Tagawa, M. (2001) *Cancer Gene Ther.* **8**, 414–420.
8. Dan, S., Tsunoda, T., Kitahara, O., Yanagawa, R., Zembutsu, H., Katagiri, T., Yamazaki, K., Nakamura, Y. & Yamori, T. (2002) *Cancer Res.* **62**, 1139–1147.
9. Cho, R. J., Campbell, M. J., Winzeler, E. A., Steinmetz, L., Conway, A., Wodicka, L., Wolfsberg, T. G., Gabrielian, A. E., Landsman, D., Lockhart, D. J. & Davis, R. W. (1998) *Mol. Cell* **2**, 65–73.
10. Hughes, T. R., Marton, M. J., Jones, A. R., Roberts, C. J., Stoughton, R., Armour, C. D., Bennett, H. A., Coffey, E., Dai, H., He, Y. D., *et al.* (2000) *Cell* **102**, 109–126.
11. Hill, W. E. (1996) *Crit. Rev. Food Sci. Nutr.* **36**, 23–173.
12. Holodniy, M. (1994) *Clin. Lab. Med.* **14**, 335–349.
13. Grayson, T. H., Cooper, L. F., Wrathmell, A. B., Roper, J., Evenden, A. J. & Gilpin, M. L. (2002) *Immunology* **106**, 273–283.
14. Parker, R. M. & Barnes, N. M. (1999) *Methods Mol. Biol.* **106**, 247–283.
15. Hod, Y. (1992) *BioTechniques* **13**, 852–854.
16. Lockhart, D. J., Dong, H., Byrne, M. C., Follettie, M. T., Gallo, M. V., Chee, M. S., Mittmann, M., Wang, C., Kobayashi, M., Horton, H. & Brown, E. L. (1996) *Nat. Biotechnol.* **14**, 1675–1680.
17. Livak, K. J., Flood, S. J., Marmaro, J., Giusti, W. & Deetz, K. (1995) *PCR Methods Appl.* **4**, 357–362.
18. Becker-Andre, M. & Hahlbrock, K. (1989) *Nucleic Acids Res.* **17**, 9437–9446.
19. McCulloch, R. K., Choong, C. S. & Hurley, D. M. (1995) *PCR Methods Appl.* **4**, 219–226.
20. Freeman, W. M., Walker, S. J. & Vrana, K. E. (1999) *BioTechniques* **26**, 112–125.
21. Bustin, S. A. (2000) *J. Mol. Endocrinol.* **25**, 169–193.
22. Hayward-Lester, A., Oefner, P. J. & Doris, P. A. (1996) *BioTechniques* **20**, 250–257.
23. Zhang, J., Day, I. N. & Byrne, C. D. (2002) *Nucleic Acids Res.* **30**, e20.
24. Yan, H., Yuan, W., Velculescu, V. E., Vogelstein, B. & Kinzler, K. W. (2002) *Science* **297**, 1143.
25. Velculescu, V. E., Zhang, L., Vogelstein, B. & Kinzler, K. W. (1995) *Science* **270**, 484–487.