Plant Mol Biol Rep (2009) 27:315–320
DOI 10.1007/s11105-009-0089-0

Arbitrary Multi-gene Reference for Normalization of Real-Time PCR Gene Expression Data

Helen H. Tai · Gregory Conn · Charlotte Davidson · H. W. (Bud) Platt

Published online: 13 February 2009

Abstract Analysis of gene expression using real-time reverse transcription polymerase chain reaction (RT-PCR) requires reference genes to normalize expression values between samples. We have developed a novel reference for real-time RT-PCR using an arbitrary primer to amplify a random set of genes. The arbitrary primer amplifies over 30 genes, whose cumulative expression as measured by real-time RT-PCR closely follows that of *UBQ 11*, an *Arabidopsis thaliana* gene that is used as a reference on microarrays. The expression of arbitrary genes is also compared with potato (*Solanum tuberosum* spp. *tubersosum*) housekeeping genes and was shown to be stable during *Phytophthora infestans* infection.

Keywords Arbitrary primers · Gene expression · Real-time PCR · Reference genes

Abbreviations

DNA	deoxyribonucleic acid
MoMLV	Moloney murine leukemia virus
RNA	ribonucleic acid
RT-PCR	reverse transcription polymerase chain reaction
DDRT-PCR	differential display reverse-transcription polymerase chain reaction

H. H. Tai (✉) · G. Conn · C. Davidson
Agriculture and Agri-Food Canada, Potato Research Centre,
P.O. Box 20280, 850 Lincoln Rd.,
Fredericton, NB E3B 4Z7, Canada
e-mail: taih@agr.gc.ca

H. W. (Bud) Platt
Agriculture and Agri-food Canada,
Crops and Livestock Research Centre,
440 University Ave.,
Charlottetown, PEI C1A 4N6, Canada

Introduction

Gene expression analyses are playing an increasingly important role in understanding of biological phenomena. Large amounts of data on differential gene expression are being generated through expressed sequence tag libraries, microarrays, serial analysis of gene expression, complementary DNA (cDNA)-amplified fragment length polymorphism, and differential display reverse transcription polymerase chain reaction (DDRT-PCR) among other techniques. Validation of differential gene expression discovered using genome-wide technologies is important, and real-time RT-PCR is being applied for this purpose (Dallas et al. 2005; Provenzano and Mocellin 2007; Wang et al. 2006). Real-time RT-PCR can provide a wide dynamic range of quantification and high levels of accuracy while using relatively lower amounts of RNA as template (Bustin et al. 2005; Czechowski et al. 2004; Higuchi et al. 1992). Experimental and technical variation needs to be controlled in real-time PCR and normalization strategies using housekeeping genes as references have been employed (Radonic et al. 2004; Reid et al. 2006). However, in many cases, the expression of a single housekeeping gene is not consistent and cannot be used for normalization. Universal reference systems combining data from several housekeeping genes (Bower et al. 2007) or addition of spike-in control RNAs into samples have been used as alternatives (Smith et al. 2003). However, the use of several reference genes represents extra cost that is significant when large numbers of samples are involved. Furthermore, spike-in controls do not account for variability in RNA isolation or quality of the RNA sample.

Arbitrary primers have been used for a range of applications including DDRT-PCR (Liang and Pardee 1992; McClelland et al. 1995) and as genetic markers for

mapping and DNA fingerprinting (Caetano-Anolles 1993). DDRT-PCR involves the amplification of PCR products using an arbitrary and an anchored oligo d(T) primer; however, RT-PCR products can be amplified with arbitrary primers alone (McClelland et al. 1995; Tai et al. 2007). A typical DDRT-PCR fingerprint can yield 50–100 bands on a gel, and differences in banding patterns are indicative of differential gene expression between samples. However, the majority of genes are not differentially expressed. In plants, only 5–10% of genes are tissue-specific (Galbraith and Birnbaum 2006). Therefore, the combined yield of all arbitrary RT-PCR products mostly represents genes that do change in expression and has a potential for use as a reference for quantitative RT-PCR. In this study, a reference method for real-time RT-PCR employing arbitrary primers is developed in *Arabidopsis thaliana* and *Solanum tuberosum*.

Materials and Methods

Preparation of Plant Tissues

Sterile *A. thaliana* seeds were sown on Murashige and Skoog media containing 1.5% agar on top of nylon filters. *A. thaliana* seeds were imbibed for 0, 1, 2, and 3 days of imbibition at 4°C and 3 days of imbibition at 4°C plus 1 day at 25°C (day 4).

Tetraploid *S. tuberosum* ssp. *tuberosum* L. cv. Green Mountain and Kathadin were grown in pots in the greenhouse. To induce biotic stress, plants were spray-inoculated with a suspension of *Pythophthora infestans* sporangia. The experiment was replicated with three plants for each cultivar. Leaves were harvested before inoculation and at 24 h after inoculation.

Extraction of RNA

Total RNA was prepared from 500 mg of plant tissue from imbibed *A. thaliana* seeds and *S. tuberosum* leaves using a hot borate extraction method (Wan and Wilkins 1994). The yield and quantity of RNA were determined by absorbance spectra between 220 and 320 nm.

Reverse Transcription Polymerase Chain Reaction

Total RNA was pretreated with RNAse-free DNAse (Invitrogen, Burlington, Canada), and 1 μg was reverse-transcribed using Moloney murine leukemia virus (MoMLV) reverse transcriptase (Invitrogen, Burlington, Canada) in a 20-μl reaction mix according to manufacturer's instruction. Messenger RNA of 0.5 μg was reverse-transcribed using MoMLV reverse transcriptase (Invitrogen, Carlsbad, USA), 40°C for 60 min in a 20-μl reaction mix with an oligo $d(T)_{11}$ primer.

Arbitrary RT-PCR amplification of 1 μl of cDNA fragments from 2-day imbibed *A. thaliana* seeds was done using 1 U Platinum Taq polymerase (Invitrogen, Mississauga, Canada.) using the following arbitrary primer: 5′ GGTACTCCAC. Cycling conditions using four different annealing temperatures were tested, 39°C, 40°C, 41°C, and 42°C. Cycling conditions were as follows: 2 min initial denaturation at 95°C; 35 cycles of 95°C 30 s, 39 (40, 41, or 42)°C 30 s, 72°C 30 s; final elongation 72°C 10 min; and final hold 6°C. Negative control PCR amplifications were done with DNAse-treated total RNA without reverse transcription and no DNA template. PCR products were

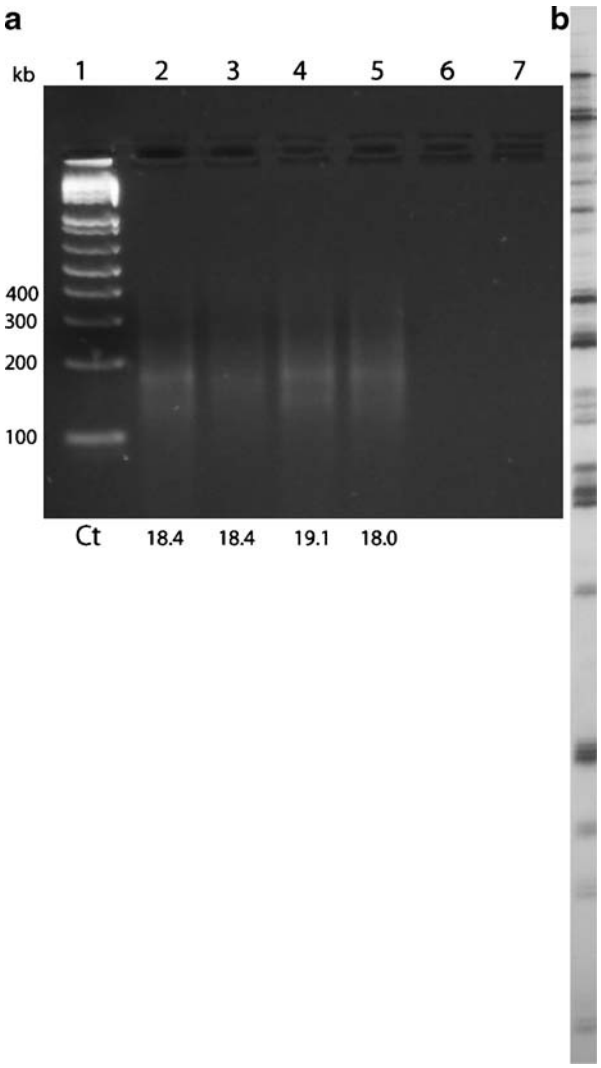

Fig. 1 *Arabidopsis thaliana* RT-PCR products with the arbitrary primer amplified using different melting temperature. **a** Two percent agarose gel electrophoresis. *Lane 1* marker, *lane 2* 39°C, *lane 3* 40°C, *lane 4* 41°C, *lane 5* 42°C, *lane 6* RT negative control, *lane 7* no DNA negative control. **b** 6% denaturing PAGE of arbitrary primed RT-PCR products

separated using gel electrophoresis with 2% agarose gels in Tris–acetate–EDTA buffer. Gels were stained with ethidium bromide and digitally photographed using the Alpha-DigiDoc gel documentation system.

Arbitrary PCR amplification of cDNA fragments was also done in the presence of ^{35}S dATP (3,000 Ci/mmol) under the same cycling conditions. Amplified products were separated on a 6% denaturing polyacrylamide gel electrophoresis. The gel was transferred to Whatman 1 filter paper and dried at 80°C using a vacuum gel dryer. The dried gel was exposed to autoradiographic film with intensifying screens at −80°C overnight.

Real-Time Reverse Transcription Polymerase Chain Reaction

Real-time PCR was done using the Dynamo kit (New England Biolabs, Mississauga, Canada) and 1 μl of cDNA and 1.8 μM of arbitrary primers according to manufacturer's instructions. The Dynamo kit uses the fluorescent dye, SYBR green, to detect PCR products. The ABI 7500 real-time thermocycler (Applied Biosystems, Foster City, USA) was used. The cycle threshold (Ct) value is the number of cycles required to accumulate enough SYBR green fluorescent signal to exceed the threshold (background) level. The Ct value is proportional to the amount of RT-PCR product and was used for quantification. Cycling conditions for real-time RT-PCR were the same as described above for the arbitrary primer. Unless otherwise indicated, the melting temperature used for arbitrary real-time RT-PCR was 40°C. Data acquisition was at 72°C on the SYBR channel. Real-time RT-PCR primers and conditions for the *A. thaliana UBQ 11* was previously described (Tai et al. 2005). The potato e1fα, β-tubulin, ribosomal protein L2, and cyclophilin real-time RT-PCR was done according to Nicot et al. (2005), and PR-1b1 was done according to Evers et al. (2006). Relative quantification was used to measure amounts of arbitrary and *UBQ 11* RT-PCR products from *A. thaliana*. For relative quantification, oligo d(T)-primed cDNA from *A. thaliana* dry seed total RNA was serially diluted (1, 1/2, 1/4, 1/8, and 0) to generate a standard curve of Ct values. The standard curve was used to quantify amounts of *UBQ 11* and arbitrary RT-PCR products from seeds imbibed for 1, 2, 3, and 4 days. The gene expression values obtained are the amount of RT-PCR product in the sample relative to dry seed. Three replicates were done for each sample.

Absolution quantification was used to quantify real-time RT-PCR results for arbitrary genes, ef1α, β-tubulin, ribosomal protein L2, cyclophilin, and PR-1b1. For absolute quantification standards, DNA generated by RT-PCR was reamplifed by PCR to generate larger quantities for generation of standard curves. Quantification of DNA for standard curve generation was done by UV spectrophotometry. Three replicates were done for each sample.

Results and Discussion

In *A. thaliana*, only 5–10% of the genes are tissue-specific (Galbraith and Birnbaum, 2006), indicating that most expressed genes do not show high variation in transcript levels. Results from DDRT-PCR, which uses arbitrary primers, also demonstrate a low level of differentially displayed bands (Liang and Pardee 1992). Quantification of DDRT-PCR data has shown that the percentage of fungal genes showing three- to ninefold increased expression after phytoalexin treatment is only 0.03% and that approximately the same percentage of genes show decreased expression (Venkatesh et al. 2005). Results were obtained in an *A. thaliana* DDRT-PCR study examining circadian patterns of gene expression, which found that only 0.017% of the

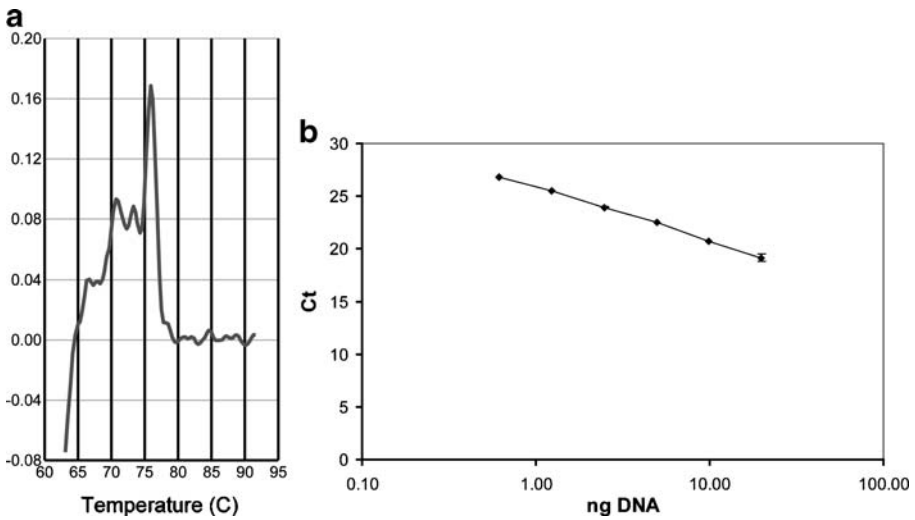

Fig. 2 Arbitrary real-time RT PCR characteristics. **a** Melting curve for arbitrary real-time RT-PCR melting curve. The template used was 2-day-imbibed *Arabidopsis thaliana* seeds. b Linear dose response of arbitrary RT-PCR to dilutions of *Arabidopsis* cDNA template. Real-time PCR was done on dilutions of cDNA from 2-day-imbibed *Arabidopsis thaliana* seeds from 0.5 to 20 ng using arbitrary primers. The Ct value for each dilution is plotted against the nanogram of cDNA in each dilution. DNA concentrations were determined by UV spectrophotometry

bands were showing circadian cycling. In a black spruce DDRT-PCR study examining seed germination, 2.2% of the 5,000 bands were differentially displayed between dry seed and imbibed seed (Tai et al. 2007). These studies provide evidence that arbitrary RT-PCR mostly amplifies genes that do not change in expression.

The reference method for real-time RT-PCR developed in this study employs an arbitrary primer to amplify multiple random genes. PCR with arbitrary primers was carried out at a low annealing temperature to allow for mispriming to increase the number of genes amplified. SYBR green dye was used to quantify DNA so the fluorescence signal is derived from multiple genes. If the number of random genes amplified by an arbitrary primer is high, then the effect of large increases in expression of a single gene will be offset by another gene that is decreased in expression. With a large enough random sample of genes, arbitrary real-time RT-PCR Ct values can represent the steady-state level of transcripts. In this case, arbitrary real-time RT-PCR can be an effective reference for data normalization that is akin to global normalization that is used for microarrays (Quackenbush 2002).

Arbitrary primer concentrations for real-time RT-PCR were tested, and 1.8 μM was found to be optimal (data not shown). In addition, low melting temperature is used with the arbitrary RT-PCR to allow for annealing to multiple targets to increase the number of bands that are amplified. Real-time RT-PCR using the arbitrary primer and *A. thaliana* total RNA from 3-day imbibed seeds was done, and products were run on agarose gel electrophoresis. Figure 1a shows that the arbitrary RT-PCR products run as a smear, indicating that multiple genes were amplified. The effect of annealing temperature during PCR cycling on RT-PCR products amplified was also tested. At 39°C, 41°C, and 42°C, preferential amplification of a band at 180 bp could be detected. At 40°C, preferential amplification of a single product was minimal. The Ct values for each annealing temperature did now show a lot of variation indicating that similar total amounts of RT-PCR products were amplified. Amplification at 40°C annealing temperature was used for the rest of the study. Arbitrary RT-PCR was also done in the presence of radioactively labeled ^{35}S [dATP], and amplified products were run on a higher resolution 6% polyacrylamide gel electrophoresis to better visualize multiple PCR products (Fig. 1b). As many as 45 bands can be resolved. Other arbitrary primers were tested but did not produce over 30 bands, which was set as the minimum number required (data not shown).

Figure 2a shows the melting curve for arbitrary real-time RT-PCR products. Multiple peaks are present, and the data acquisition temperature for real-time PCR was set at 72°C. Figure 2b demonstrates a linear relationship between Ct values and dilutions of cDNA from imbibed *A. thaliana* seeds. These results demonstrate that arbitrary real-time RT-PCR can be used within the range of cDNA concentrations that are typically used.

Polyubiquitin genes are used as reference genes for normalization of gene expression in *A. thaliana* (Czechowski et al. 2005). *UBQ 11* gene expression is stable during seed imbibition (Tai et al. 2005) and is used as a reference gene on the Affymetrix AG *A. thaliana* gene chip (www.affymetrix.com). Arbitrary real-time RT-PCR was compared with *UBQ 11* real-time RT-PCR in this study. Serial dilutions of cDNA reverse-transcribed from dry seed total RNA were used as a standard for relative quantification used in this experiment. *A. thaliana* seeds at 1, 2, 3, and 4 days of imbibition were selected for examination, since

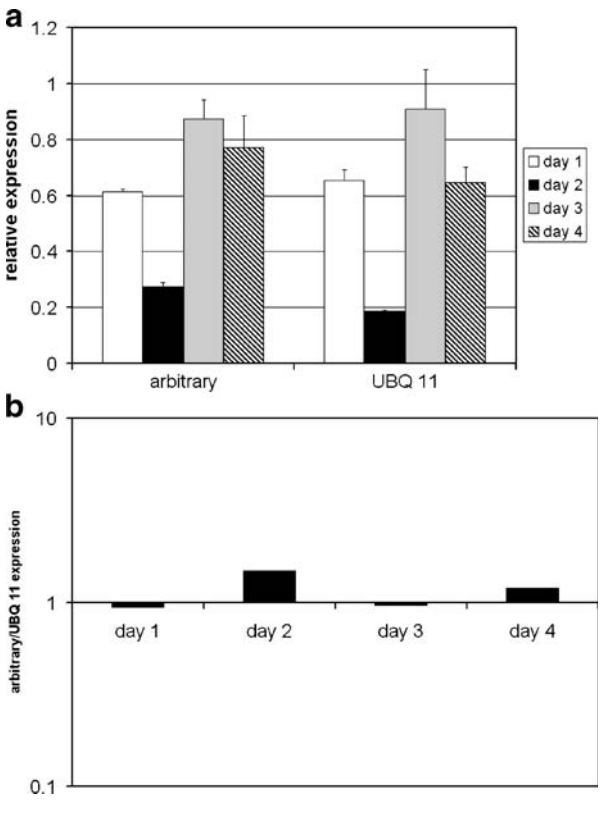

Fig. 3 Expression of arbitrary genes is similar to the housekeeping gene, *UBQ 11*, during *Arabidopsis thaliana* seed imbibition. Real-time RT-PCR results are quantified relative to varying amounts of *Arabidopsis thaliana* dry seed RNA. Two-fold serial dilutions (1, 1/2, 1/4, 1/8, and 0) of cDNA synthesized from dry seed RNA were used to generate the standard curve for real-time PCR. **a** cDNA synthesized from *Arabidopsis thaliana* seeds imbibed for 1, 2, 3, and 4 days at 4°C were analyzed for expression of *UBQ 11* and arbitrary genes using real-time PCR. Expression of arbitrary genes and *UBQ 11* at day 1 (*white bars*), day 2 (*black bars*), day 3 (*gray bars*), and day 4 (*striped bars*) of seed imbibition is shown. Relative expression is the amount of *Arabidopsis thaliana* dry seed cDNA that would be required to give the same amount of RT-PCR product. The mean of three replicate real-time RT-PCR reactions are presented. The *error bars* represent standard error of the mean. **b** Graph of average relative expression values for arbitrary real-time RT-PCR are divided by average values for UBQ 11

changes in patterns of gene expression have been observed during this period of development (Gallardo et al. 2001; Nakabayashi et al. 2005; Tai et al. 2005). The data show that the relative expression values for arbitrary and *UBQ 11* genes are in the same range (Fig. 3a). Real-time RT-PCR relative expression values for arbitrary genes were divided by values for *UBQ 11* (arbitrary/*UBQ 11*) to demonstrate that the differences between the two genes is similar for the imbibed seed samples (Fig. 3b). The lowest arbitrary/*UBQ 11* value is 0.94 (day 3) and the highest is 1.48 (day 2). Expression of arbitrary genes and *UBQ 11* was lower at 2 days after imbibition compared with the rest of the time points, which is likely due to variation in the RNA quality rather than decreased expression of arbitrary genes or *UBQ 11*.

The expression of arbitrary genes were also examined in *S. tuberosum*. Results are presented in Fig. 4. Absolute quantification was used to determine the nanogram amounts of each real-time RT-PCR product. Arbitrary RT-PCR results in the high amounts of PCR products amplified (Fig. 4a) compared with the single gene RT-PCR amplification (Fig. 4b–f). The PR-1b1 gene was shown by others to be upregulated following *Phytophthora infestans* inoculation in *Solanum phureja* (Evers et al. 2006). Figure 4f shows that PR-1b1 is upregulated in *P. infestans* inoculated *S. tuberosum* cv. Green Mountain and to a lesser extent in cv. Katahdin. The results from PR-1b1 are used to demonstrate that changes in gene expression patterns that have taken place are induced in response to the *P. infestans* inoculation. The expression of the housekeeping genes ef1α, β-tubulin, ribosomal protein L2, and cyclophilin are presented in Fig. 4b–e, respectively. Previous studies done by others (Nicot et al. 2005) have demonstrated that ef1α is the most stable of these housekeeping genes during *Phythophthora infestans* infection, which is confirmed by the results presented in this paper. The expression of housekeeping genes β-tubulin, ribosomal protein L2, and cyclophilin all show significant fluctuation between uninoculated and inoculated samples, whereas ef1α expression does not differ significantly. The real-time RT-PCR results using the arbitrary primer show that there is also stability in the amount of arbitrary RT-PCR products between uninoculated and inoculated samples (Fig. 4a). The results show that the

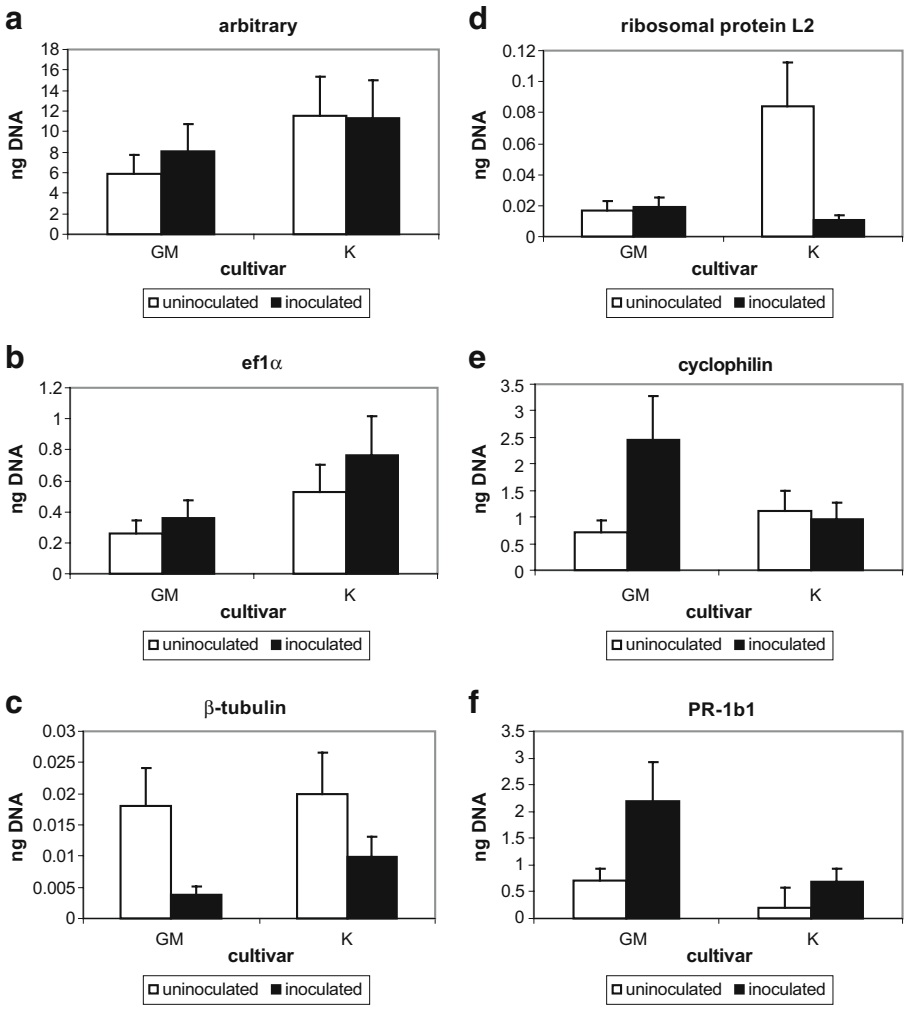

Fig. 4 Expression of arbitrary genes, housekeeping genes and PR-1b1 in potato. Absolute quantification of real-time RT-PCR results are presented as ng DNA of amplified product from uninoculated (*white bars*) and *Phytophthora infestans* inoculated (*black bars*) leaves from Green Mountain (GM) and Katahdin (K) potato varieties. **a** arbitrary genes, **b** ef1α, **c** β-tubulin, **d** ribosomal protein L2, **e** cyclophilin, and **f** PR-1b1. The mean of three replicate RT-PCR reactions for three plants (nine readings) are presented. The *error bars* represent standard error of the mean

arbitrary RT-PCR products have higher stability than the housekeeping genes.

The amount of arbitrary RT-PCR products was found to be higher in cv. Katahdin than in cv. Green Mountain. The expression of ef1α was also higher for cv. Katahdin over cv. Green Mountain. These results may be due to differences in the quality of RNA between cv. Katahdin and cv. Green Mountain.

This study has demonstrated that arbitrary primers can be used in real-time RT-PCR. The real-time RT-PCR results obtained are a pooled signal from multiple random genes. This signal is stable through developmentally induced changes in gene expression patterns in *A. thaliana* and with biotic stress in *S. tuberosum*. Based on the results from this study, we propose the use of arbitrary real-time RT-PCR as a reference for normalization of real-time RT-PCR data. Arbitrary real-time RT-PCR was done successfully with two different plant species, which indicates that it may have application in multiple species.

Acknowledgments The authors would like to thank Ian Macdonald who provided technical assistance for this study.

References

Bower NI, Moser RJ, Hill JR, Lehnert SA (2007) Universal reference method for real-time PCR gene expression analysis of preimplantation embryos. Biotechniques 42:199–206 doi:10.2144/000112314

Bustin SA, Benes V, Nolan T, Pfaffl MW (2005) Quantitative real-time RT-PCR—a perspective. J Mol Endocrinol 34:597–601 doi:10.1677/jme.1.01755

Caetano-Anolles G (1993) Amplifying DNA with arbitrary oligonucleotide primers. PCR Methods Appl 3:85–94

Czechowski T, Bari RP, Stitt M, Scheible WR, Udvardi MK (2004) Real-time RT-PCR profiling of over 1400 *Arabidopsis* transcription factors: unprecedented sensitivity reveals novel root- and shoot-specific genes. Plant J 38:366–379 doi:10.1111/j.1365-313X.2004.02051.x

Czechowski T, Stitt M, Altmann T, Udvardi MK, Scheible WR (2005) Genome-wide identification and testing of superior reference genes for transcript normalization in *Arabidopsis*. Plant Physiol 139:5–17 doi:10.1104/pp.105.063743

Dallas PB, Gottardo NG, Firth MJ, Beesley AH, Hoffmann K, Terry PA et al (2005) Gene expression levels assessed by oligonucleotide microarray analysis and quantitative real-time RT-PCR—how well do they correlate? BMC Genomics 6:59 doi:10.1186/1471-2164-6-59

Evers D, Schweitzer C, Nicot N, Gigliotti S, Herrera MR, Hausman JF et al (2006) Two PR-1 loci detected in the native cultivated potato *Solanum phureja* appear differentially expressed upon challenge by late blight. Physiol Mol Plant Pathol 67:155–163 doi:10.1016/j.pmpp.2005.12.003

Galbraith DW, Birnbaum K (2006) Global studies of cell type-specific gene expression in plants. Annu Rev Plant Biol 57:451–475 doi:10.1146/annurev.arplant.57.032905.105302

Gallardo K, Job C, Groot SPC, Puype M, Demol H, Vandekerckhove J et al (2001) Proteomic analysis of *Arabidopsis* seed germination and priming. Plant Physiol 126:835–848 doi:10.1104/pp.126.2.835

Higuchi R, Dollinger G, Walsh PS, Griffith R (1992) Simultaneous amplification and detection of specific DNA sequences. Biotechnology (NY) 10:413–417 doi:10.1038/nbt0492-413

Liang P, Pardee AB (1992) Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. Science 257:967–971 doi:10.1126/science.1354393

McClelland M, Mathieu-Daude F, Welsh J (1995) RNA fingerprinting and differential display using arbitrarily primed PCR. Trends Genet 11:242–246 doi:10.1016/S0168-9525(00)89058-7

Nakabayashi K, Okamoto M, Koshiba T, Kamiya Y, Nambara E (2005) Genome-wide profiling of stored mRNA in *Arabidopsis thaliana* seed germination: epigenetic and genetic regulation of transcription in seed. Plant J 41:697–709 doi:10.1111/j.1365-313X.2005.02337.x

Nicot N, Hausman JF, Hoffmann L, Evers D (2005) Housekeeping gene selection for real-time RT-PCR normalization in potato during biotic and abiotic stress. J Exp Bot 56:2907–2914 doi:10.1093/jxb/eri285

Provenzano M, Mocellin S (2007) Complementary techniques: validation of gene expression data by quantitative real time PCR. Adv Exp Med Biol 593:66–73 doi:10.1007/978-0-387-39978-2_7

Quackenbush J (2002) Microarray data normalization and transformation. Nat Genet 32(Suppl):496–501 doi:10.1038/ng1032

Radonic A, Thulke S, Mackay IM, Landt O, Siegert W, Nitsche A (2004) Guideline to reference gene selection for quantitative real-time PCR. Biochem Biophys Res Commun 313:856 doi:10.1016/j.bbrc.2003.11.177

Reid K, Olsson N, Schlosser J, Peng F, Lund S (2006) An optimized grapevine RNA isolation procedure and statistical determination of reference genes for real-time RT-PCR during berry development. BMC Plant Biol 6:27 doi:10.1186/1471-2229-6-27

Smith RD, Brown B, Ikonomi P, Schechter AN (2003) Exogenous reference RNA for normalization of real-time quantitative PCR. Biotechniques 34:88–91

Tai HH, Tai GCC, Beardmore T (2005) Dynamic histone acetylation of late embryonic genes during seed germination. Plant Mol Biol 59:909–925 doi:10.1007/s11103-005-2081-x

Tai HH, Williams M, Iyengar A, Yeates J, Beardmore T (2007) Regulation of the β-hydroxyacyl ACP dehydratase gene of *Picea mariana* by alternative splicing. Plant Cell Rep V26:105

Venkatesh B, Hettwer U, Koopmann B, Karlovsky P (2005) Conversion of cDNA differential display results (DDRT-PCR) into quantitative transcription profiles. BMC Genomics 6:51 doi:10.1186/1471-2164-6-51

Wan CY, Wilkins TA (1994) A modified hot borate method significantly enhances the yield of high-quality RNA from cotton (*Gossypium hirsutum* L.). Anal Biochem 223:7–12 doi:10.1006/abio.1994.1538

Wang Y, Barbacioru C, Hyland F, Xiao W, Hunkapiller KL, Blake J et al (2006) Large scale real-time PCR validation on gene expression measurements from two commercial long-oligonucleotide microarrays. BMC Genomics 7:59 doi:10.1186/1471-2164-7-59