

Multiplex cDNA quantification method that facilitates the standardization of gene expression data

Osamu Gotoh¹, Yasufumi Murakami² and Akira Suyama^{1,*}

¹Department of Life Sciences and Institute of Physics, Graduate School of Arts and Sciences, The University of Tokyo, 3-8-1 Komaba, Meguro-ku, Tokyo 153-8902 and ²Department of Biological Science and Technology, Graduate School of Industrial Science and Technology, Tokyo University of Science, 2641 Yamazaki, Noda, Chiba 278-8510, Japan

Received October 10, 2010; Revised February 15, 2011; Accepted February 24, 2011

ABSTRACT

Microarray-based gene expression measurement is one of the major methods for transcriptome analysis. However, current microarray data are substantially affected by microarray platforms and RNA references because of the microarray method can provide merely the relative amounts of gene expression levels. Therefore, valid comparisons of the microarray data require standardized platforms, internal and/or external controls and complicated normalizations. These requirements impose limitations on the extensive comparison of gene expression data. Here, we report an effective approach to removing the unfavorable limitations by measuring the absolute amounts of gene expression levels on common DNA microarrays. We have developed a multiplex cDNA quantification method called GEP-DEAN (Gene expression profiling by DCN-encoding-based analysis). The method was validated by using chemically synthesized DNA strands of known quantities and cDNA samples prepared from mouse liver, demonstrating that the absolute amounts of cDNA strands were successfully measured with a sensitivity of 18 zmol in a highly multiplexed manner in 7 h.

INTRODUCTION

Methods for gene expression analysis are increasingly important for biological and medical research (1–6). The DNA microarray-based method is now one of the most widely used methods for gene expression analysis. This method is highly parallel, so that thousands of genes can be analyzed all at once. However, the observed data are

substantially affected by microarray platforms and RNA references, because this method provides only relative expression values (7–9). Therefore, the standardization of gene expression profiling (GEP) data are a subject of considerable interest in the microarray community.

The MicroArray Quality Control Project and the External RNA Control Consortium are organized efforts to resolve the GEP-data standardization issue (10–12). They have developed standardization methods by use of two types of RNA reference materials: assay process references and universal hybridization references. Although this method is suited to determining relative expression differences between samples, it would be highly desirable to have a method which could measure the absolute amounts of target gene transcripts, such as mRNA copies per cell. This type of measurement is specifically important for sharing interplatform gene expression data in public repositories (7,8), as well as for conducting systems biology research (13).

For measurement of the absolute amounts of nucleic acids, a quantitative PCR (qPCR)-based method is widely used as a validation method. This method is highly sensitive for a low quantity of samples. However, assay designs and protocols must be considerably adjusted in order to achieve accurate quantifications using this method (14). Furthermore, determining the absolute expression levels for many target genes requires a large number of reactions using serially diluted templates in order to make standard curves (14,15). Therefore, determining the absolute amounts using the qPCR method is so time and cost intensive that it is not appropriate for large-scale gene expression profiling.

Under these circumstances, a new method is required for both highly parallel and sensitive quantification of absolute amounts of transcripts. Here, we report a novel gene expression profiling method to determine the absolute amount of cDNA (the copy number of cDNA

*To whom correspondence should be addressed. Tel/Fax: +81 3 5454 6528; Email: suyama@dna.c.u-tokyo.ac.jp

strands) for many target genes in a highly parallel and sensitive manner. The method is called GEP-DEAN because the gene expression profiling is performed using the DEAN (DCN-encoding-based analysis) technology (16), which is a technique for analyzing target information by means of well-designed DNA-tag sequences called DNA-Coded Numbers (DCNs), originally developed for reliable DNA computing (17). The use of DCNs is highly advantageous in that gene expression profiling of different sets of target genes can be performed by using DNA microarrays with the same set of DNA probe sequences. Even an analysis for single nucleotide polymorphism typing can be performed by using the same DNA microarray and basically the similar protocol to gene expression profiling (18,19).

Currently, the GEP-DEAN method can successfully measure the absolute amount of target cDNA with a sensitivity of 18 zmol (approximately 10 000 copies) and multiplicity of about 300 target genes in 7 h. A validation using cDNA samples prepared from mouse liver revealed that the method accurately quantified cDNA samples equivalent to 18 ng total RNA.

MATERIALS AND METHODS

Synthetic DNA samples

In order to investigate the dependence of the Cy5/Cy3 ratio on DNA quantities, we prepared synthetic DNA samples that contained 291 kinds of 30-base DNA strands at various known quantities. Their sequences were parts of the cDNA strands of *Cyanidioschyzon merolae* (red algae). These DNA strands were commercially synthesized and purified by reverse-phase cartridge chromatography (Operon). For comparison of GEP-DEAN and qPCR, we prepared a mixture of 57 kinds of 99-base DNA strands in various known quantities. Their sequences are portions of yeast cDNA strands. These DNA strands were commercially synthesized and purified by polyacrylamide gel electrophoresis (Sigma Genosys). The concentration of every DNA strand was determined by UV-absorbance measurements at 260 nm (20). The sequences used in this study are listed in the Supplementary Tables 1–3.

The cDNA samples

Complementary DNA samples were prepared from total RNA samples extracted from liver tissues obtained from acetaminophen administered or no drug-administered 8-week-old male BALB/c mice (Clea Japan). All animal procedures were carried out in compliance with the institutional animal ethics guidelines. The protocols of the animal procedures were as follows: mice were intraperitoneally injected with 367 mg/kg acetaminophen. The injected dose was determined by the concentration of LD₅₀ given in the material safety data sheet. The mice were sacrificed at 6 h after dosing, followed by perfusion with 0.1 M phosphate buffered saline (PBS). The liver samples were stored in RNAlater buffer (Ambion) with 1 ml solution per 0.1 g liver tissue at 4°C. After removing the RNAlater, the samples were subjected to dissection

and homogenization in TRIzol reagent (Invitrogen). The total RNA samples were extracted using a TRIzol kit according to the manufacturer's protocols. Contaminated DNA in total RNA samples was digested with DNase I and the following purification was performed using an RNeasy mini kit (Qiagen) according to the manufacturer's protocol. The RNA quality was checked by UV-absorbance measurements. The total RNA samples were reverse-transcribed with oligo-dT primers of SuperScript III (Invitrogen) according to the manufacturer's protocol. The prepared cDNA samples were stored at -20°C until use.

Synthetic reference DNA and internal DNA

Reference DNA and internal DNA were commercially synthesized and purified by reverse-phase cartridge chromatography or polyacrylamide gel electrophoresis (Operon, Sigma Genosys). The concentration of every DNA strand was determined by UV-absorbance measurements at 260 nm (20). The sequences used in this study are listed in the Supplementary Tables 1–3.

DCNs

DCNs are 92-base DNA-tag sequences that are composed of four sections, designated as SD, D1_{*j*1} (*j*1 = 0, 1, 2, ..., N_{D1} - 1), D2_{*j*2} (*j*2 = 0, 1, 2, ..., N_{D2} - 1) and ED (Figure 1b). The characteristics of DCNs were described previously (18). Briefly, these sequences are designed to have a uniform length, a uniform melting temperature and have no potential for mishybridizations or stable self-folded structures. The sequences are listed in the Supplementary Table 4.

Probe sequences

Probe sequences of 30 bases in length were designed so as to be complementary to target specific sequences (TSS). They were designed according to the previously published criteria (21). Each probe sequence was divided into two parts: a 15-base 5'-probe sequence (PS1_{*i*}) and 15-base 3'-probe sequence (PS2_{*i*}) (Figure 1b). The probe pairs were filtered to minimize the mispriming risk of molecular translation table (MTT) strands, which could result in false PCR products in the encoding step. The probe pairs used in this study are listed in the Supplementary Tables 1–3.

MTT

The 38-base 5'-MTT strands (CS-PS1_{*i*}; *i* = 0, 1, 2, ..., N_T - 1) were commercially synthesized and purified by reverse-phase cartridge chromatography (Operon, Hokkaido System Science). The 3'-MTT strands were prepared as follows: the 38-base SD-attached PS2_{*i*} strands (PS2_{*i*}-SD; *i* = 0, 1, 2, ..., N_T - 1) and 92-base 3'-amino-modified complementary DCN strands (cDCN_{*j*}; *j* = 0, 1, 2, ..., N_T - 1) were commercially synthesized and purified by reverse-phase cartridge chromatography (Sigma Genosys, Operon, Hokkaido System Science). The cDCN-templated extensions of PS2_{*i*}-SD from the 3'-end of SD were performed in individual

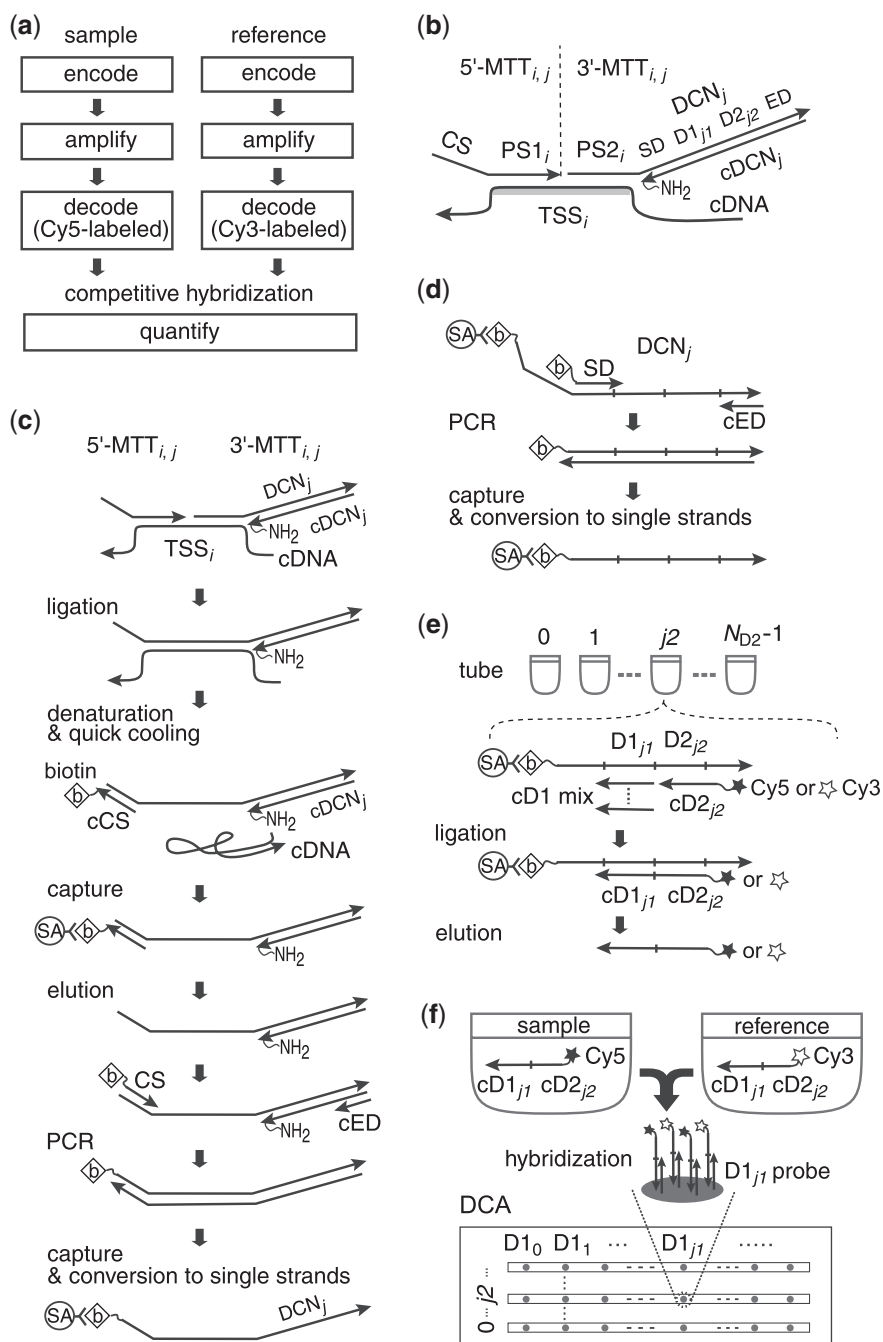


Figure 1. Schematic representation of GEP-DEAN. (a) Outline of the assay procedures. (b) Structure of molecular translation table (MTT) that connects the target cDNA with a target specific sequence TSS_i to the corresponding two-digit DNA-coded number DCN_j composed of the common start digit (SD), the first digit D1_{j1}, the second digit D2_{j2} and the common end digit (ED). (c) Encoding step. The encoding reaction for the target cDNA strands with TSS_i starts with ligation of 5'-MTT_{i,j} and 3'-MTT_{i,j} strands that hybridize adjacently with TSS_i. After ligation of the MTT strands, the reaction mixture is subjected to incubation at high temperature and a subsequent quick cooling to dissociate the target cDNA strands and to make 3'-biotinylated cCS (cCS-b) hybridize. The ligation products with cCS-b bound are captured by streptavidin-coupled magnetic beads (SA-beads) and then excess free MTT strands and the dissociated target cDNA strands are washed out. After elution of the ligation products, a primer pair of 5'-b-CS and cED strands is added to the reaction mixture and the ligation products are amplified by PCR to further remove the free 3'-MTT strands. The amplified ligation products are then captured by SA-beads and converted to single strands by alkali wash. (d) Amplification step. The DCN region of the single-stranded ligation products is amplified by PCR with a primer pair of 5'-biotinylated SD and cED. The amplified products are then captured by SA-beads and converted to single strands by alkali wash. (e) Decoding step. The two-digit DCNs are decoded with respect to the second digit (D2). The solution of the single-stranded amplified products of the sample and that of the reference are divided into N_{D2} tubes. To the j₂-th tube, a fluorescence-labeled (Cy5- and Cy3-labeled for sample and reference, respectively) cD2_{j2} strand and the mixture of all cD1 strands are added. When a DCN_j strand corresponding to the target cDNA with TSS_i is present, a cD2_{j2} and a cD1_{j1} strand are joined together by DCN-templated ligation to produce a decoding product cD2_{j2}-cD1_{j1}. The decoding products of the sample and reference are thus labeled with different fluorescence colors. (f) Quantification step. The decoding products of the sample and those of the reference in the j₂-th tube are competitively hybridized in the j₂-th capillary of a DNA capillary array (DCA), which is a kind of DNA microarray composed of an array of capillaries with DNA probes immobilized on their inner surface. All capillaries have a common set of DNA probes D1_{j1} (j₁ = 0, 1, 2, ..., N_{D1} - 1). The absolute amount of the target cDNA with TSS_i is obtained from the Cy5/Cy3 fluorescence intensity ratio measured for a spot of D1_{j1} probe in the j₂-th capillary.

wells of a 96-well PCR plate. Each well contained a unique combination of PS2_r-SD and cDCN_j. The reactions were performed in 40 μl reaction cocktails containing 0.5 μM PS2_r-SD and 0.5 μM cDCN_j, 1 mM MgSO₄, 1× *Thermococcus kodakaraensis* (KOD) DNA polymerase buffer and 0.02 U/μl KOD-plus-DNA polymerase (Toyobo) using a thermal cycler (DNA Engine PTC-200, Bio-Rad). The mixtures were first incubated at 94°C for 2 min, followed by five cycles of 94°C for 15 s, 64°C for 30 s and 68°C for 20 s. Note that no extension of cDCN_j strands from the 3'-end occurs because their 3'-ends are amino-modified. The uniquely assigned 3'-MTT_{i,j} strands were mixed into one tube for purification by a Wizard SV Gel and PCR Clean-Up System (Promega), following the manufacturer's protocol. The purified products were checked by capillary electrophoresis (Agilent 2100 Bioanalyzer, Agilent) and phosphorylated at 37°C for 30 min in 70 μl reaction cocktails containing 1 mM ATP and 0.43 U/μl T4 polynucleotide kinase (Takara) and terminated at 70°C for 5 min. The prepared 5'-MTT and 3'-MTT strands were stored at -20°C until use. The assignments of DCNs to the probe sequences used in this study are given in the Supplementary Tables 1–3.

Encoding

A sample and reference were assayed in parallel. Each reaction was first held at 95°C for 3 min in a 29.5 μl reaction mixture containing 10 fmol of each MTT set (5'-MTT and 3'-MTT) and 1× *Taq* DNA ligase buffer (New England BioLabs), followed by annealing at 0.1°C/s until reaching 45°C. Then, 0.5 μl of *Taq* DNA ligase (New England BioLabs) was added to each reaction mixture. Ligation reactions were performed at 45°C for 1 h. After the ligation reactions were completed, the products were captured by addition of 10 pmol of 3'-biotinylated cCS (cCS-b) and incubation at 95°C for 3 min, followed by quick cooling at 4°C. The reaction mixtures were captured by adding 0.3 mg of streptavidin-coupled magnetic beads (Dynabeads MyOne streptavidin C1, Invitrogen's Dynal) and incubating at room temperature for 15 min, followed by washing twice at room temperature with 1× binding and washing buffer (B&W: 1 M NaCl, TE, pH 8.0), according to the manufacturer's protocol. The captured products were suspended in 50 μl of distilled water and then eluted with 30 μl after incubation at 95°C for 3 min. The eluted products were amplified in 50 μl reaction mixtures containing 0.4 μM 5'-biotinylated CS (b-CS), 1 μM cED, 0.2 mM dNTP mixture, 1 mM MgSO₄, 1× KOD DNA polymerase buffer and 0.02 U/μl KOD-plus-DNA polymerase. The PCR reactions were first held at 94°C for 2 min, followed by 20 cycles of 94°C for 15 s, 64°C for 30 s and 68°C for 10 s and the reaction was stopped at 4°C. Each PCR product was captured by addition of 0.3 mg of streptavidin-coupled magnetic beads and incubation at room temperature for 15 min, followed by washing at room temperature with 1× B&W. Then, an alkali wash was performed using 1× alkali buffer (0.1 M NaOH, 0.05 M NaCl). Finally, the products were washed twice with 1× B&W.

Amplification

The encoding products were suspended in 100 μl distilled water. Then, 10 μl of each suspension was amplified with 0.4 μM 5'-biotin-modified SD (b-SD) and 1 μM cED, following the PCR protocol at the encoding step. The amplified products were captured by addition of 0.3 mg of streptavidin-coupled magnetic beads, followed by washing with 1× alkali buffer and 1× B&W, according to the washing protocol at the encoding step.

Decoding

The fluorescently labeled strands (cD2–cD1 with Cy5 or Cy3 modification) were produced by DCN-templated ligation using 100 kinds of complementary D1s (cD1_{j1}, $j1 = 0, 1, 2, \dots, 99$) with 5'-phosphorylation (0.1 μM each) and complementary D2 (1 μM) with Cy5 (sample) or Cy3 (reference) modification of the 5'-end. The reaction was performed at 50°C for 15 min in 50 μl reaction mixtures including 1× *Taq* ligase buffer and 0.4 U/μl *Taq* DNA ligase. Then, the ligation products were washed with 1× B&W twice and suspended in 50 μl distilled water. The suspensions were incubated at 95°C for 3 min and then eluted with 40 μl. An 18.25 μl aliquot of the eluted products of the sample and an 18.25 μl aliquot of the eluted products of the reference were mixed in the final 50 μl buffer containing 5× SSC and 0.2% SDS and hybridized at 50°C for 30 min on DNA capillary arrays that were spotted with 100 kinds of D1s. The hybridized array was washed at 50°C for 15 min with the washing buffer containing 0.1% SDS and 0.1× SSC, then scanned at photomultiplier voltages of 600 V for Cy5 and Cy3 using a commercially available DNA microarray scanner (GenePix 4000B unit, Axon Instruments).

Quantification

Fluorescence images were analyzed by the commercially available software package GenePix Pro 5.1 (Axon Instruments). The local background-subtracted median intensities of Cy5 and Cy3 were used in further calculations. A signal-to-noise ratio (SNR) of 3 was employed for the cut-off value of the signal intensities of Cy5 and Cy3. The SNR was defined as follows: (signal–background)/(standard deviation of background). The linear and nonlinear regression analyses were performed using the R environment (available at <http://www.r-project.org>).

Quantitative PCR

The qPCR was performed in a 30 μl cocktail containing 1× KOD DNA polymerase buffer, 0.2 mM dNTPs, 1 mM MgSO₄, 0.02 U/μl KOD-plus-DNA polymerase, 1× SYBR Green I (Invitrogen), 0.3 μM each of the forward and reverse primers and template DNA using 96-well plates and adhesive seals on a DNA Engine Opticon 2 system (Bio-Rad). The template DNA was either a sample DNA mixture or serial dilutions (0.5–500 amol/reaction) of standard DNA. The temperature profile of the reaction was as follows: 94°C for 2 min, followed by 30 cycles of 94°C for 15 s, 64°C for 30 s, 68°C for 30 s and a plate reading. Each reaction was completed with a

melting curve analysis to check the specificity of amplification. The primer sets used in this experiment are listed in the Supplementary Table. Quantification cycle (Cq) (22) values were calculated using an Opticon Monitor 3 real-time PCR machine (Bio-Rad). The absolute amounts of target DNA strands were determined using the standard curves obtained from the relationship between the concentrations of standard DNA templates and the corresponding Cq values. A standard curve was constructed for every target DNA. The analysis was performed using the qpcR package for the R environment (23).

RESULTS

GEP-DEAN method

GEP-DEAN is a DEAN method for gene expression profiling. By GEP-DEAN, highly parallel cDNA quantification is made for any set of target genes with common DNA microarrays. The method involves the following four steps: encoding, amplification, decoding and quantification (Figure 1a). A sample containing cDNA strands and a reference containing a known concentration of chemically synthesized DNA strands with target cDNA-specific sequences (TSSs) are treated in parallel to determine the absolute amounts of cDNA strands in the sample.

The encoding step converts each of the target DNA strands into the corresponding DCN strand through DNA-templated ligation of a molecular translation table (MTT) composed of 5'-MTT and 3'-MTT strands (Figure 1b). The 5'-MTT_{*ij*} and 3'-MTT_{*ij*} strands for the target DNA with TSS_{*i*} have the 5'-probe sequence PS1_{*i*} complementary to the 3'-half of TSS_{*i*} and the 3'-probe sequence PS2_{*i*} complementary to the 5'-half of TSS_{*i*}, respectively. The 5'-end portion of 5'-MTT_{*ij*} has the common capture sequence CS that is used to capture encoded products on streptavidin-coupled magnetic beads (SA-beads) with a biotinylated strand, cCS-b, i.e. complementary to CS. The use of the common capture strand cCS-b instead of biotinylated 5'-MTT strands is more cost-effective. The 3'-end portion of 3'-MTT_{*ij*}, on the other hand, has a two-digit DNA-coded number, DCN_{*j*}, corresponding to the target cDNA with TSS_{*i*}. DCN_{*j*} consists of four sections: SD, D1_{*j1*}, D2_{*j2*} and ED running in a 5'-3' direction. The start digit (SD) and the end digit (ED) are common among all DCNs. D1_{*j1*} and D2_{*j2*} (*j1* = 0, 1, 2, ..., *N*_{D1} - 1; *j2* = 0, 1, 2, ..., *N*_{D2} - 1) stand for the first and the second digit of the two-digit DCN, respectively. To create one-to-one mappings between the target cDNA and DCN, *N*_{D1}*N*_{D2} must not be less than the total number of target genes *N*_T. The use of the two-digit DCN allows a large scale gene expression profiling to be performed with small scale DNA microarrays. An expression profile of 1000 target genes, e.g. can be obtained with 10 U of 100-probe common DNA microarrays. The DNA sequences of SD, D1, D2 and ED are chosen from the improved set of orthonormal DNA sequences originally developed for reliable DNA computing (17).

The encoding reaction starts with ligation of a 5'-MTT_{*ij*} and a 3'-MTT_{*ij*} strand sitting adjacently on TSS_{*i*}

(Figure 1c). The target DNA strands are then removed by denaturing and quickly cooling, which helps the cDNA strands form self-folded structures stimulating strand separation. The ligation products bound to cCS-b are captured by SA-beads and eluted after washing-out excess free MTT strands as well as the separated target DNA strands. The isolated ligation products are not yet ready for the amplification step, because they are often contaminated with minute amounts of free 3'-MTT strands, which increases the level of background noise in quantification. To further remove the free 3'-MTT strands, the ligation products are amplified by PCR with a common primer pair: 5'-b-CS and complementary ED (cED) primer, with which no free 3'-MTT strands are amplified. Note that no extension of cDCN_{*j*} strands from the 3'-end occurs during the amplification because their 3'-ends are amino-modified. The amplified ligation products are then isolated by SA-beads to proceed to the amplification step.

The amplification step is to amplify the DCN region of the ligation products with another common primer pair: 5'-biotinylated SD and cED primer (Figure 1d). The amplified DCN double-strands are captured by SA-beads and then made single-stranded to obtain the DCN-amplified products for the decoding step.

The decoding step is to decode the two-digit DCNs with respect to the second digit D2 (Figure 1e). The solution of the DCN-amplified products is divided into *N*_{D2} tubes. To the *j2*-th tube, a fluorescence-labeled cD2_{*j2*} strand and the mixture of all cD1 (cD1_{*j1*}; *j1* = 0, 1, 2, ..., *N*_{D1} - 1) strands are added. When a DCN_{*j*} strand corresponding to the target cDNA with TSS_{*i*} is present, a cD2_{*j2*} and a cD1_{*j1*} strand are joined together to produce a decoding product cD2_{*j2*}-cD1_{*j1*}. The DCN-amplified products of the sample are decoded with Cy5-labeled cD2 strands. Those of the reference, on the other hand, are decoded with Cy3-labeled cD2 strands. Decoding products of the sample and reference are thus labeled with different fluorescence colors.

The final quantification step is to measure the fluorescence intensity ratios of the Cy5-labeled decoding products (sample) to the Cy3-labeled ones (reference) to obtain a gene expression profile (Figure 1f). The decoding products of the sample and those of the reference in the *j2*-th (*j2* = 0, 1, 2, ..., *N*_{D2} - 1) tube are competitively hybridized in the *j2*-th capillary of a DNA capillary array (DCA) with a common set of DNA probes D1_{*j1*} (*j1* = 0, 1, 2, ..., *N*_{D1} - 1). DCA is a kind of DNA microarray composed of an array of capillaries with DNA probes immobilized on their inner surface.

The Cy5 and Cy3 fluorescence intensities measured from the spot of the D1_{*j1*} probe give a Cy5/Cy3 ratio, *r*(*j*), which is proportional to the ratio *r*_{DEC}(*j*) of the decoding product quantity for a sample to that for a reference. The quantity *S*_{DEC}(*j*) of a decoding product cD2_{*j2*}-cD1_{*j1*} derived from target DNA with TSS_{*i*} in a sample is given by $\varepsilon_{\text{DEC}}(j) \cdot \varepsilon_{\text{AMP}} \cdot \varepsilon_{\text{ENC}}(i, j) \cdot S(i)$, where $\varepsilon_{\text{ENC}}(i, j)$, ε_{AMP} and $\varepsilon_{\text{DEC}}(j)$ are the efficiency of the encoding, the amplification and the decoding step and *S*(*i*) is the quantity of target DNA with TSS_{*i*} in the sample. Here, $\varepsilon_{\text{ENC}}(i, j)$ was found not to be constant despite an uniform melting temperature of TSS_{*i*},

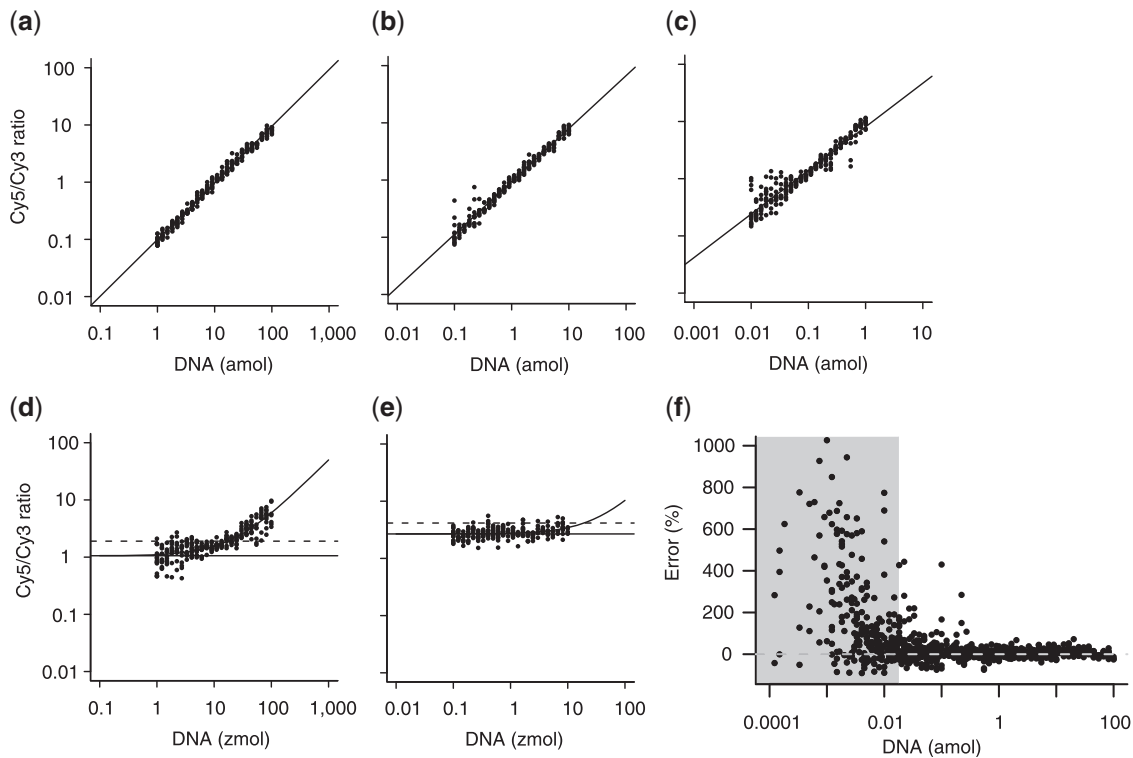


Figure 2. Dependence of the Cy5/Cy3 ratio on DNA quantity. The quantities (concentrations) of DNA strands in the serial dilution samples were (a) 1–100 amol (0.033–3.3 pM), (b) 0.1–10 amol (3.3 fM–0.33 pM), (c) 0.01–1 amol (0.33 to 33 fM), (d) 1–100 zmol (33 aM–3.3 fM) and (e) 0.1–10 zmol (3.3 aM–0.33 fM). Those of DNA strands in the reference mixtures were uniform: (a) 10 amol (0.33 pM), (b) 1 amol (33 fM), (c) 0.1 amol (3.3 fM), (d) 10 zmol (0.33 fM) and (e) 1 zmol (33 aM). The slopes of the regression lines in a, b and c were 0.99, 0.93 and 0.76, respectively. The R^2 -values were 0.99, 0.98 and 0.92, respectively. The solid curves in d and e are the ones fitted to the model in Equation (2). The horizontal solid lines represent background levels. The dashed lines represent a signal-to-noise ratio of 3. (f) The quantification error in a–e. The shaded area is below the lower limit of detection.

However, for every TSS_{*i*}, $\varepsilon_{\text{ENC}}(i, j)$ was not affected by $S(i)$ (Supplementary Data 1). The ε_{AMP} depended on neither the sequence nor the quantity of DCN strands, because the DCN strands made of the orthonormal sequences were uniformly amplified with the common primer pair SD and cED. $\varepsilon_{\text{DEC}}(j)$ was slightly dependent on the DNA sequence of DCN, even though D1 and D2 are orthonormal sequences, because the decoding step involves a ligation reaction as in the case of the encoding step. Similarly, the quantity $R_{\text{DEC}}(j)$ of a decoding product cD2_{*j*1}-cD1_{*j*2} derived from synthetic DNA with TSS_{*i*} in a reference is given by $\varepsilon_{\text{DEC}}(j) \cdot \varepsilon_{\text{AMP}} \cdot \varepsilon_{\text{ENC}}(i, j) \cdot R(i)$, where $R(i)$ is the quantity of synthetic DNA with TSS_{*i*} in the reference. The ratio $r_{\text{DEC}}(j)$ is thus $S_{\text{DEC}}(i)/R_{\text{DEC}}(i) = S(i)/R(i)$. Since the Cy5/Cy3 ratio $r(j)$ is proportional to $r_{\text{DEC}}(j)$, the quantity $S(i)$ can be determined from the Cy5/Cy3 ratio as

$$S(i) = \alpha r(j) R_c. \quad (1)$$

when a mixture of synthetic DNA strands of a uniform quantity R_c is used for the reference. The proportionality constant α can be one by adjusting the sensitivity of the microarray scanner to Cy5 and Cy3 fluorescence.

Dependence of the Cy5/Cy3 ratio on DNA quantity

Equation (1) indicates the proportionality between the Cy5/Cy3 ratio and the quantity of DNA in a sample.

We have investigated this linear relationship using 10-fold serial dilutions of a DNA mixture that contained 291 kinds of chemically synthesized 30-base DNA strands in various known quantities.

As the quantity of sample DNA was reduced from 100 amol (3.3 pM) to 0.1 amol (3.3 fM), the Cy5/Cy3 ratio decreased proportionally (Figure 2a and b). The slope of the linear regression and the R^2 -value for the log–log plot in Figure 2a were 0.99 and 0.99, respectively, which means that the Cy5/Cy3 ratio was proportional to $S(i)^{0.99}$ and 99% of the variance in the Cy5/Cy3 ratio could be explained by the proportional relation. In Figure 2b, the slope and the R^2 -value were slightly decreased to 0.93 and 0.98, respectively, still indicating the proportional relation.

As the DNA quantity was further reduced, however, the Cy5/Cy3 ratio was no longer proportional to the DNA quantity. The slope of linear regression for the log–log plot was decreased to 0.76 in the range of 0.01–1 amol DNA (Figure 2c). In the range of 1–100 zmol and 0.1–10 zmol DNA, the data no longer followed the linear relation (Figure 2d and e). After an elaborate investigation, it turned out that the proportionality vanished due to the contamination of the encoding product with a minute amount of free 3'-MTT. Even after substantially reducing the relative quantity of free 3'-MTT by PCR amplification of the ligation product, the encoding product eluted from SA-beads was still contaminated with a non-negligible

amount of free 3'-MTT, which was amplified and decoded together with the encoding product to increase the background noise.

The Cy5/Cy3 ratio $r(j)$ has a different expression when taking the contamination into account. $S_{DEC}(j)$ is given by $\varepsilon_{DEC}(j) \varepsilon_{AMP} \{ \varepsilon_{ENC}(i,j) S(i) + \delta \}$, where δ stands for the quantity of free 3'-MTT strands with DCN_j present in the encoding product. The factor, δ is independent of DCN as long as each 3'-MTT strand has a uniform concentration. Similarly, $R_{DEC}(j)$ is given by $\varepsilon_{DEC}(j) \varepsilon_{AMP} \{ \varepsilon_{ENC}(i,j) R_c + \delta \}$. Therefore, the Cy5/Cy3 ratio $r(j)$ can be expressed by

$$r(j) = \frac{1 S(i) + \delta / \varepsilon_{ENC}(i,j)}{\alpha R_c + \delta / \varepsilon_{ENC}(i,j)} \quad (2)$$

The data in Figure 2d and e were well fitted to the model in Equation (2) with a constant $\delta / \varepsilon_{ENC}(i,j)$.

Multiple DNA quantification by GEP-DEAN

The expression used to determine the quantity $S(i)$ of DNA with TSS_i in a sample from the observed Cy5/Cy3 ratio $r(j)$ can be derived from Equation (2):

$$S(i) = \alpha r(j) R_c + \{ \alpha r(j) - 1 \} \delta / \varepsilon_{ENC}(i,j). \quad (3)$$

When the amount of free 3'-MTT in the encoding product is negligible, the expression becomes equal to Equation (1). Otherwise, it has a non-negligible second term containing unknown factors, δ and $\varepsilon_{ENC}(i,j)$. We have determined the values of unknown factors and the proportionality constant α by using the standard curve obtained from the Cy5/Cy3 ratios of DNA strands of known quantities spiked in a sample. The standard curve was fitted to the model by substituting an average encoding efficiency $\overline{\varepsilon_{ENC}}$ for $\varepsilon_{ENC}(i,j)$ in Equation (3) to determine the unknown factors.

Figure 2f is the result of quantification of the serial dilution DNA samples shown in Figure 2a–e. For each of the dilution samples, 27 DNA strands out of the 291 kinds of DNA strands were used as the spiked-in DNA to draw the standard curve. Another 264 kinds of DNA strands in each sample were quantified and then all of the quantification data were compiled into a single graph in Figure 2f. The quantification error, which was defined as (observed quantity – actual quantity)/actual quantity, was within $\pm 20\%$ for 78% of DNA samples and $\pm 30\%$ for 90% of DNA samples of >0.1 amol, respectively. The lower limit of detection, which was defined as the DNA quantity with the Cy5/Cy3 ratio deviating from the background noise by a factor of three, was 18 zmol (0.6 fM, approximately 10 000 copies) (Figure 2d and e). The quantification error increased substantially below the lower limit of detection (shaded area in Figure 2f). The pattern of assignments of DCNs to target DNAs did not affect the results of quantification (Supplementary Data 2).

Comparison with the qPCR method

The GEP-DEAN method was compared with the qPCR method, which is the gold standard method for measuring

absolute amounts of DNA. The DNA sample employed for the comparison was a mixture of 57 kinds of chemically synthesized 99-base DNA strands in various known quantities (concentrations) ranging from 1 to 100 amol (0.033–3.3 pM), which were determined by using UV-absorbance at 260 nm and dilution factors.

First, we compared the reproducibilities of the two methods. Independent duplicate analyses demonstrated that the measurement of the absolute amount of DNA by GEP-DEAN was as highly reproducible as that by qPCR (Figure 3a and b). The correlation plot of the duplicate measurement by GEP-DEAN exhibited a regression line with a slope of one and R^2 -value of 1.00. For the measurement by qPCR, the correlation plot indicated a regression line with a slope of one and an R^2 -value of 0.97.

Then, we compared the accuracies of the two methods. Based on the degree of agreement between the measured DNA quantity and the actual DNA quantity determined using UV-absorbance at 260 nm and dilution factors, GEP-DEAN was more accurate than qPCR (Figure 3c and d). In addition, while only one kind of DNA in a sample can be quantified by a single qPCR analysis, numerous kinds of DNA in a sample can be quantified in parallel by a single GEP-DEAN analysis. Therefore, the measurement of the absolute amount of DNA by GEP-DEAN is more efficient than that by qPCR.

Gene expression profiling by GEP-DEAN

Parallel quantifications of cDNA prepared from mouse liver cells were performed to examine the reproducibility, accuracy and sensitivity of gene expression profiling by GEP-DEAN. cDNA strands for 273 mouse genes (Supplementary Table) with various expression levels were quantified in parallel by using Equation (3), in which $\varepsilon_{ENC}(i,j)$ was replaced with $\overline{\varepsilon_{ENC}}$. The values of α , δ and $\overline{\varepsilon_{ENC}}$ were determined using the standard curves constructed with synthetic 30-base DNA strands of known quantities added to the samples. The standard curves were constructed for each D2 value because the value of α varied with the tube in which the decoding step was performed (Figure 4b).

The reproducibility of cDNA quantification by GEP-DEAN was confirmed by using a sample containing a quantity of cDNA equivalent to 2 μ g of total RNA. Duplicate measurements exhibited such a high reproducibility that the log–log plot had a regression line with a slope of one and an R^2 -value of 0.98 (Figure 4a).

The accuracy of cDNA quantification by GEP-DEAN was examined by using serial dilutions of the original cDNA sample measured in Figure 4a. Since the accuracy is defined as the degree of agreement between a measured quantity and the actual quantity, the measured values should be compared with the actual values. However, in contrast to chemically synthesized DNA, there are no objective means to obtain the actual value of the absolute amount of cDNA in a sample. We therefore, considered the value measured from the original cDNA sample and dilution factors as the actual value in a dilution sample in order to examine the accuracy of cDNA quantification. The absolute amounts of cDNA measured for the serial

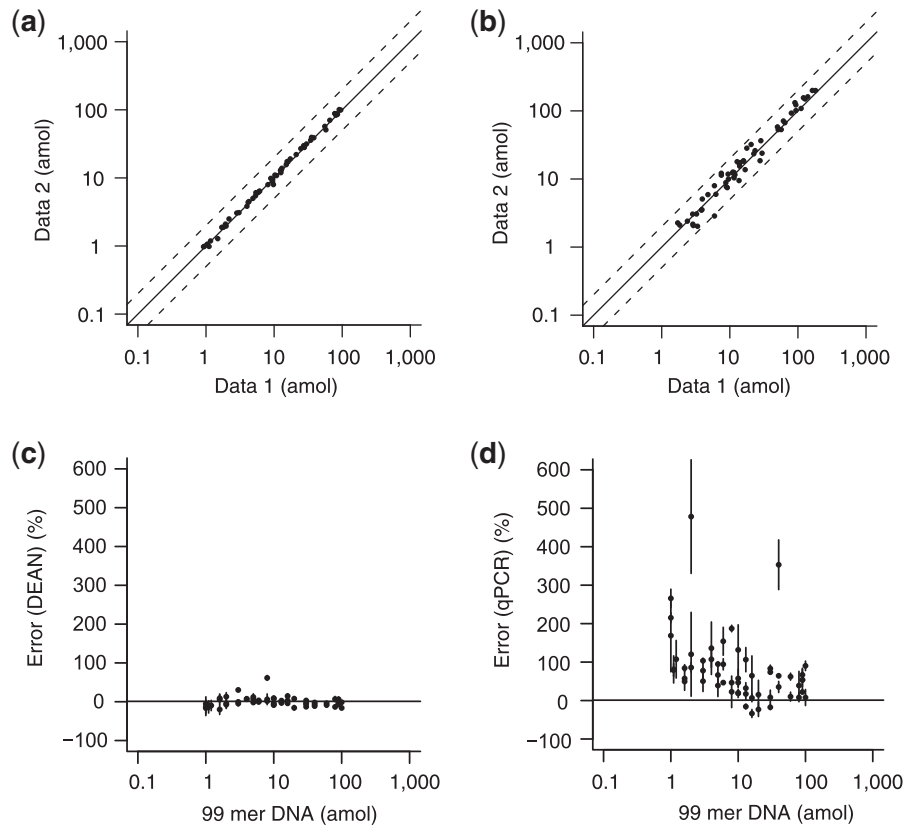


Figure 3. Comparison between GEP-DEAN and qPCR. The reproducibility of GEP-DEAN (a) and that of qPCR (b) are plotted. The solid lines show the ideal lines at an angle of 45°. The dashed lines are 2-fold or 0.5-fold of the ideal lines. The R^2 -values on a log-log scale were 1.00 and 0.97, respectively. The accuracies of GEP-DEAN (c) and qPCR (d) are plotted. The horizontal lines show the level of no quantification error. Error bars show the SDs of replicated measurements.

dilution samples were successfully decreased in proportion to dilution factors, so that they were consistent with the estimated actual values (Figure 4c-f). In the log-log plots for the dilution samples equivalent to 1, 0.5, 0.25 and 0.125 μg of total RNA, the data were fitted to a regression line with a slope of one. Their R^2 -values were 0.98, 0.95, 0.89 and 0.77, respectively.

The minimum sample quantity needed for gene expression profiling by GEP-DEAN can be estimated from the expression profiling data in Figure 4. The expression profiles in Figure 4 show that many genes were expressed at a level over 1 amol in a cDNA sample equivalent to 1 μg total RNA. Considering that the sensitivity of GEP-DEAN determined by the experiments using chemically synthesized DNA in Figure 2 was 18 zmol, the minimum sample quantity can be estimated to be a cDNA sample equivalent to 18 ng total RNA.

Analysis of differentially expressed genes

By using the gene expression profiles of mouse liver cells treated with acetaminophen (APAP), we have demonstrated that differentially expressed genes could be directly determined from measurements of the absolute amount of cDNA by GEP-DEAN without common controls or any additional normalizations. APAP is one of the most

commonly used drugs for pain and fever and also an important cause of serious liver injury (24).

For determination of genes differentially expressed due to APAP administration, the ratio of the absolute cDNA quantity in an APAP-administered sample to that in a no drug-administered sample was calculated from the observed cDNA quantities. The ratio was then compared with the quantity ratio expected for genes with unchanged expression levels. The significance of the difference between the two ratios was examined by the two-sided Welch's t -test. In testing the null hypothesis that the expression level of the i -th gene is not affected by APAP administration, the following statistic was employed:

$$t = \frac{\bar{X}_i - \bar{X}_U}{\sqrt{V_i/N_i + V_U/N_U}}, \quad (4)$$

where \bar{X}_i , V_i and N_i are the mean, variance and size of the ratios for the i -th tested gene and \bar{X}_U , V_U and N_U are those of the ratios for unchanged genes. The degree of freedom ϕ is given by:

$$\frac{1}{\phi} = \frac{c^2}{N_i - 1} + \frac{(1 - c)^2}{N_U - 1} \quad c = \frac{V_i}{N_i} / \left(\frac{V_i}{N_i} + \frac{V_U}{N_U} \right). \quad (5)$$

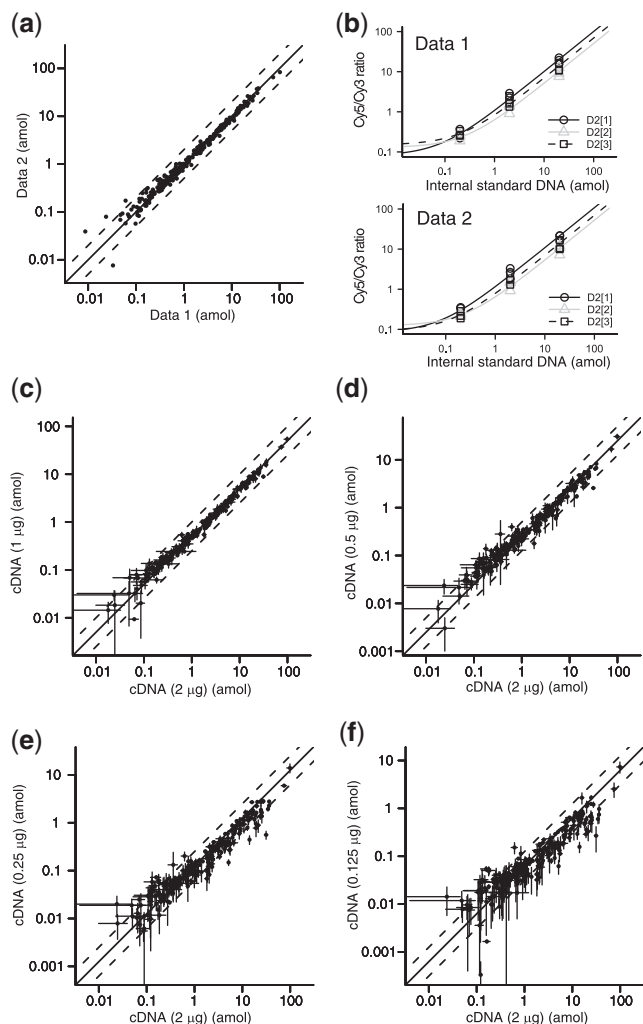


Figure 4. Parallel quantifications of cDNA prepared from mouse liver cells. (a) Reproducibility of two independent quantifications of cDNA equivalent to 2 μg of total RNA. The slope and R^2 -value of the regression line were 1.00 and 0.98 on a log-log scale. (b) Standard curves used in a. (c–f) Quantification of cDNA equivalent to 2 μg of total RNA was compared to that of serially dilution samples containing a quantity of cDNA equivalent to 1 (c), 0.5 (d), 0.25 (e) and 0.125 μg (f) of total RNA. The solid lines at an angle of 45° are the expected lines calculated from the quantifications of cDNA equivalent to 2 μg of total RNA and the dilution factors. The dashed lines are 2- or 0.5-fold of the solid lines. The R^2 -values of the regression lines with a slope of one were 0.98, 0.95, 0.89 and 0.77, respectively. Error bars show the SDs of triplicate measurements.

Figure 5 shows a plot of \bar{X}_i calculated from quadruplicate measurements of the cDNA quantities in the APAP administered (after 6 h) and no drug-administered samples against the mean of the absolute cDNA quantity of the i -th gene in a no drug-administered sample. When the amount of total cDNA is equal in all samples, \bar{X}_U is expected to be one. However, as shown in Figure 5, many genes whose expression levels were expected to be unchanged had a ratio of around 0.82, because the total amount of cDNA prepared from the same amount of total RNA varied among the samples. In the actual comparison of gene expression profiles, the information about the total

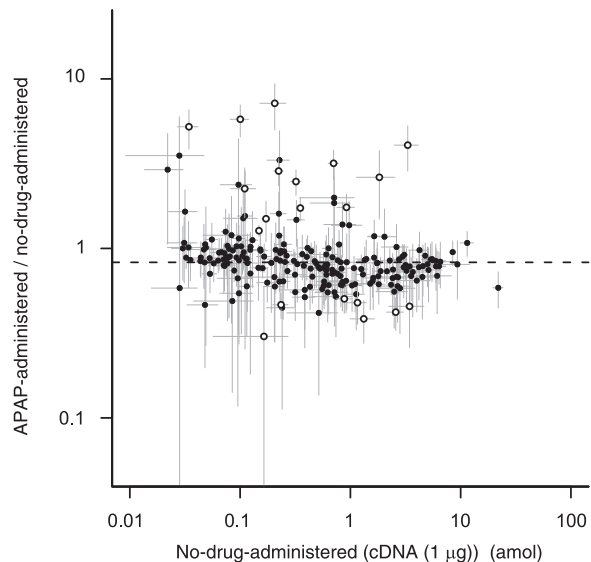


Figure 5. Genes expressed differentially in the APAP-administered mouse liver. The Y-axis represents the ratio of the absolute amounts of cDNA in an APAP-administered sample to those in a no-drug-administered one. The X-axis represents the absolute amounts of cDNA in a no drug-administered sample. The dashed line represents the median ratio value. Error bars on both axes show the SDs. The candidates for significantly changed genes (open circle) were determined by a two-sided Welch's t -test ($P < 0.05$) using quadruplicate data.

amount of cDNA is usually unavailable. The values for \bar{X}_U and V_U were thus determined from genes with an \bar{X}_i value close to the median ratio. These genes are safely assumed to be candidate genes with unchanged expression levels. Using \bar{X}_U and V_U calculated from the mean of \bar{X}_i and V_i of the candidate genes with \bar{X}_i of a median value $\pm 10\%$, 20 of 223 genes were found to be significantly up- or downregulated ($P < 0.05$). When the range of unchanged-gene candidates was taken as the median value $\pm 5\%$, the number of significantly regulated genes was 19 and hardly varied with the range. In this test, Bonferroni multi-test corrections were not applied because a t -test with a corrected $P < 0.00022$ ($0.05/223$) is not appropriate for screening of APAP-induced genes.

DISCUSSION

We have developed a novel gene-expression profiling method named GEP-DEAN, which enables parallel quantification of absolute cDNA amounts as low as 18 zmol (approximately 10 000 copies) with a multiplicity of 291 target genes on a 100-probe DCA in 7 h. By GEP-DEAN, a gene expression profile is obtained as the distribution of the absolute cDNA amounts of target genes. Therefore, the method can indeed provide expression data in a standardized form, allowing more extensive comparison of gene expression profiles to find differentially expressed genes without complicated data normalization. GEP-DEAN requires no sample labeling and thus results create no labeling-bias problems. Gene expression profiling of different sets of target genes can be performed

by using common DCAs with the same set of DNA probe sequences.

Technologies or approaches similar to GEP-DEAN have been reported previously, including multiple ligation-dependent probe amplification (MLPA) technology (25), cDNA-mediated annealing, selection, extension and ligation (DASL) assay (26) and padlock probes (PLPs) technology (27,28). These methods also employ ligation and PCR amplification, enabling highly sensitive, reproducible and versatile assays. However, none of these earlier methods can determine the absolute cDNA amounts.

The use of two-digit DCNs involves the decoding step, so that it results in a longer assay time than the use of single-digit DCNs. However, the two-digit DCNs allow a genome-wide GEP on small-size DNA microarrays. We have already designed about 500 kinds of orthonormal DNA sequences for D1 and D2. If 460 of these sequences are assigned to D1 and 40 of them to D2, 18 400 kinds of target genes can be analyzed all at once using DCAs with only 460 probes. Highly multiplex decoding reactions needed for a genome-wide GEP would be feasible, because the handling of liquids in this study was successfully performed using a Biomek2000 robot (Beckman Coulter). In addition, the two-digit DCNs allow a wide dynamic range of quantification. Figure 2 would demonstrate that if different D2 values are assigned for groups of genes with different expression levels, GEP can be performed with a wide dynamic range exceeding the practical dynamic range of DNA microarrays.

In Figure 3, the reproducibility and accuracy of DNA quantification by GEP-DEAN were compared with those by qPCR, which is a standard method for measuring the absolute amount of a very small quantity of DNA. Here, we briefly compare the two methods in terms of the assay cost. The comparison is made for an instance of a GEP analysis of 300 target genes, which is the analysis of the same number of target genes as that demonstrated in Figure 4. First, we compare the cost of DNA synthesis. For an assay by GEP-DEAN, 300 sets of CS-PS₁_{*i*} (38-mer), PS₂-SD (38-mer) and TSS_{*i*} (30-mer) strands should be newly synthesized for every set of target genes, whereas 300 cDCN_{*j*} (92-mer) strands used for the assay are not necessary to be synthesized for every target gene set because cDCN_{*j*} strands are used universally and at a very small quantity (1/1000 of a typical quantity of PCR primer). The other DNA strands, namely 100 cD1_{*f1*} (23-mer), 3 sets of Cy5-/Cy-3-cD2_{*f2*} (23-mer), b-SD (23-mer), cED (23-mer), cCS-b (23-mer) and b-CS (23-mer), are also used universally and at a small quantity; thus they are not necessary to be synthesized for every set of target genes either. For an assay by qPCR, on the other hand, 300 sets of DNA strands for the PCR primer pair (typically 30-mer) should be newly synthesized for every set of target genes. Therefore, the cost of the newly synthesized DNA strands for GEP-DEAN is almost 1.5× as high as that for qPCR. The GEP-DEAN assay further needs universally used DNA strands; however, their running costs are negligible compared to the cost of newly-synthesized DNA strands. Next, we compare the cost of enzymatic reactions for

PCR. For a GEP-DEAN assay, the cost of PCR is negligible because the encoding products and their DCN regions for 300 target genes are amplified all at once. For a qPCR assay, in contrast, the cost of PCR is extremely high because the number of reactions for the real time PCR is as large as 1500 (300 for sample, 1200 for construction of 300 standard curves using 4 serially diluted templates). Therefore, as for the running cost, the GEP-DEAN method is concluded to have an advantage over the qPCR method. As for the total cost including the full cost for the synthesis of universally used DNA strands, the conclusion would be highly dependent on the available cost of DNA synthesis and that of enzymes for PCR.

The GEP-DEAN method is able to analyze almost any kind of target DNA strands accurately, so that it can also be applied to alternative splicing or copy number variance analysis. Moreover, this method can be adapted for direct RNA measurements if RNA-templated ligation is employed at the encoding step. Therefore, GEP-DEAN has potential for direct simultaneous measurement of non-coding RNAs and messenger RNAs.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

FUNDING

Development of System and Technology for Advanced Measurement and Analysis (SENTAN), Japan Science and Technology Agency (to A.S.). Funding for open access charge: The University of Tokyo (to A.S.).

Conflict of interest statement. None declared.

REFERENCES

- Schena, M., Shalon, D., Davis, R.W. and Brown, P.O. (1995) Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science*, **270**, 467–470.
- 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. *et al.* (1996) Expression monitoring by hybridization to high-density oligonucleotide arrays. *Nat. Biotechnol.*, **14**, 1675–1680.
- Gresham, D., Dunham, M.J. and Botstein, D. (2008) Comparing whole genomes using DNA microarrays. *Nat. Rev. Genet.*, **9**, 291–302.
- Mardis, E.R. (2008) The impact of next-generation sequencing technology on genetics. *Trends Genet.*, **24**, 133–141.
- Tang, F., Barbacioru, C., Wang, Y., Nordman, E., Lee, C., Xu, N., Wang, X., Bodeau, J., Tuch, B.B., Siddiqui, A. *et al.* (2009) mRNA-Seq whole-transcriptome analysis of a single cell. *Nat. Methods*, **6**, 377–382.
- Geiss, G.K., Bumgarner, R.E., Birditt, B., Dahl, T., Dowidar, N., Dunaway, D.L., Fell, H.P., Ferree, S., George, R.D., Grogan, T. *et al.* (2008) Direct multiplexed measurement of gene expression with color-coded probe pairs. *Nat. Biotechnol.*, **26**, 317–325.
- Bassett, D.E. Jr, Eisen, M.B. and Boguski, M.S. (1999) Gene expression informatics—it's all in your mine. *Nat. Genet.*, **21**, 51–55.
- Brazma, A., Hingamp, P., Quackenbush, J., Sherlock, G., Spellman, P., Stoeckert, C., Aach, J., Ansorge, W., Ball, C.A., Causton, H.C. *et al.* (2001) Minimum information about a

- microarray experiment (MIAME)-toward standards for microarray data. *Nat. Genet.*, **29**, 365–371.
9. Shippy,R., Fulmer-Smentek,S., Jensen,R.V., Jones,W.D., Wolber,P.K., Johnson,C.D., Pine,P.S., Boysen,C., Guo,X., Chudin,E. *et al.* (2006) Using RNA sample titrations to assess microarray platform performance and normalization techniques. *Nat. Biotechnol.*, **24**, 1123–1131.
 10. Baker,S.C., Bauer,S.R., Beyer,R.P., Brenton,J.D., Bromley,B., Burrill,J., Causton,H., Conley,M.P., Elespuru,R., Fero,M. *et al.* (2005) The External RNA Controls Consortium: a progress report. *Nat. Methods*, **2**, 731–734.
 11. Reid,L.H., Lucas,A.B., Kopf-Sill,A.R., Chen,B., Bromley,B., Foy,C., Hinkel,C.S., Boysen,C., Liu,C.M., Ranamukha-arachchi,D. *et al.* (2005) Proposed methods for testing and selecting the ERCC external RNA controls. *BMC Genomics*, **6**, 150.
 12. Shi,L., Reid,L.H., Jones,W.D., Shippy,R., Warrington,J.A., Baker,S.C., Collins,P.J., de Longueville,F., Kawasaki,E.S., Lee,K.Y. *et al.* (2006) The MicroArray Quality Control (MAQC) project shows inter- and intraplatform reproducibility of gene expression measurements. *Nat. Biotechnol.*, **24**, 1151–1161.
 13. Cassman,M. (2005) Barriers to progress in systems biology. *Nature*, **438**, 1079.
 14. Nolan,T., Hands,R.E. and Bustin,S.A. (2006) Quantification of mRNA using real-time RT-PCR. *Nat. Protoc.*, **1**, 1559–1582.
 15. Bustin,S.A. (2000) Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays. *J. Mol. Endocrinol.*, **25**, 169–193.
 16. Nishida,N., Wakui,M., Tokunaga,K. and Suyama,A. (2001) Highly specific and quantitative gene expression profiling based on DNA computing. *Genome Inform.*, **12**, 259–260.
 17. Yoshida,H. and Suyama,A. (2000) Solution to 3-SAT by breadth first search. *DIMACS*, **54**, 9–22.
 18. Nishida,N., Tanabe,T., Hashido,K., Hirayasu,K., Takasu,M., Suyama,A. and Tokunaga,K. (2005) DigiTag assay for multiplex single nucleotide polymorphism typing with high success rate. *Anal. Biochem.*, **346**, 281–288.
 19. Nishida,N., Tanabe,T., Takasu,M., Suyama,A. and Tokunaga,K. (2007) Further development of multiplex single nucleotide polymorphism typing method, the DigiTag2 assay. *Anal. Biochem.*, **364**, 78–85.
 20. Cavaluzzi,M.J. and Borer,P.N. (2004) Revised UV extinction coefficients for nucleoside-5'-monophosphates and unpaired DNA and RNA. *Nucleic Acids Res.*, **32**, e13.
 21. Kurata,K. and Suyama,A. (1999) Probe design for DNA chips. *Genome Inform.*, **10**, 225–226.
 22. Lefever,S., Hellemans,J., Pattyn,F., Przybylski,D.R., Taylor,C., Geurts,R., Untergasser,A. and Vandesompele,J. (2009) RDML: structured language and reporting guidelines for real-time quantitative PCR data. *Nucleic Acids Res.*, **37**, 2065–2069.
 23. Ritz,C. and Spiess,A.N. (2008) qpcR: an R package for sigmoidal model selection in quantitative real-time polymerase chain reaction analysis. *Bioinformatics*, **24**, 1549–1551.
 24. Lee,W.M. (2004) Acetaminophen and the U.S. Acute Liver Failure Study Group: lowering the risks of hepatic failure. *Hepatology*, **40**, 6–9.
 25. Schouten,J.P., McElgunn,C.J., Waaijer,R., Zwijnenburg,D., Diepvens,F. and Pals,G. (2002) Relative quantification of 40 nucleic acid sequences by multiplex ligation-dependent probe amplification. *Nucleic Acids Res.*, **30**, e57.
 26. Fan,J.B., Yeakley,J.M., Bibikova,M., Chudin,E., Wickham,E., Chen,J., Doucet,D., Rigault,P., Zhang,B., Shen,R. *et al.* (2004) A versatile assay for high-throughput gene expression profiling on universal array matrices. *Genome Res.*, **14**, 878–885.
 27. Baner,J., Isaksson,A., Waldenström,E., Jarvius,J., Landegren,U. and Nilsson,M. (2003) Parallel gene analysis with allele-specific padlock probes and tag microarrays. *Nucleic Acids Res.*, **31**, e103.
 28. Ericsson,O., Jarvius,J., Schallmeiner,E., Howell,M., Nong,R.Y., Reuter,H., Hahn,M., Stenberg,J., Nilsson,M. and Landegren,U. (2008) A dual-tag microarray platform for high-performance nucleic acid and protein analyses. *Nucleic Acids Res.*, **36**, e45.