Electronic Supplementary Information (ESI)

Pyrosequencing-based barcodes for dye-free multiplex bioassay†

Zhiyao Chen,[a] Xiaoying Fu,[b] Xiaodan Zhang,[b] Xiqun Liu,[b] Bingjie Zou,[b] Haiping Wu,[b] Qinxin Song,[a] Jinheng Li,[a] Tomoharu Kajiyama,[c] Hideki Kambara[c] and Guohua Zhou*[a,b]

Table of contents

Supporting Materials and Methods of BAMGA

Reagents and chemicals

TaKaRa M-MLV RTase cDNA synthesis kit, *Mbo* I endonuclease, T4 DNA ligase, SYBR *Premix Ex* Taq^{TM}, TaKaRa M-MLV RTase cDNA Synthesis Kit and *rTaq* DNA polymerase were purchased from TaKaRa (Dalian, China). Exo$^-$ Klenow Fragment and QuantiLum recombinant luciferase were purchased from Promega (Madison, WI). TRIzol reagent, Superscript II RNase H$^-$ Reverse Transcriptase were from Invitrogen (Inc. Faraday Avenue, Carlsbad, CA), and Dynabeads M-280 Streptavidin (2.8 μm) was from Dynal Biotech ASA (Oslo, Norway). ATP sulfurylase, D-luciferin, bovine serum albumin (BSA), polyvinylpyrrolidone (PVP), adenosine 5'-phosphosulfate (APS), and apyrase were obtained from Sigma (St. Louis, MO). 2'-deoxyguanosine-5'-triphosphate (dGTP), 2'-deoxythymidine-5'-triphosphate (dTTP), 2'-deoxycytidine-5'-triphosphate (dCTP), 2', 3'-dideoxycytidine-5'-triphosphate (ddCTP), 2', 3'-dideoxyguanosine-5'- triphosphate (ddGTP), Sodium 2'-deoxyadenosine-5'-O-(1-triphosphate) (dATPαS), and Sequenase 2.0 were purchased from Amersham Pharmacia Biotech (Amersham, U.K.). BioSpin PCR purification kit was purchased from Bio Flux Corporation (Hangzhou, China). Other chemicals were of a commercially extra-pure grade. All solutions were prepared with deionized and sterilized water.

RNA extraction and ds-cDNA synthesis

Total RNAs from small intestines, right brain, large intestine, cerebellum, spleen, lung, left brain, bladder, heart, stomach, ear, and kidney tissues of a mouse were extracted using TRIzol reagent according to the manufacturer's instructions. The purity and concentration of the extracted RNA were determined by a UV-Vis spectrophotometer (NaKa Instrument Co., ltd, Japan). Then 1.5 μg of each total RNAs from different tissues was used for first-strand cDNA synthesis by Superscript II RNase H$^-$ Reverse Transcriptase, respectively. First, total RNA in RNase-free water was heated to 65 °C for 5 min and cooled on ice for at least 1 min. Then a 10 μl of reaction mixture containing 0.5 mM of each dNTP, 1×RT buffer, 1 μM oligo $(dT)_{15}$ primer, 0.5 U/μl RNase inhibitor, and 0.2 U/μl Superscript II RNase H$^-$ Reverse Transcriptase was added to each tube containing extracted RNA, and incubated at 42 °C for 60 min, followed by 93 °C for 5 min to terminate the reaction. All of the first-strand cDNA was used to synthesize double-stranded cDNA (ds-cDNA) by TaKaRa M-MLV RTase cDNA Synthesis Kit according to the manufacturer's instructions. The 50 μl of ds-cDNA synthesis reactions containing 10 μl of first-strand

cDNA solution, 10 µl of 5×ds-cDNA synthesis buffer (18.8 mM Tris-HCl, 90.6 mM KCl, 4.6 mM $MgCl_2$, 10 mM $(NH_4)_2SO_4$, pH 8.3), 0.15 mM NAD, 0.002% BSA, 3.8 mM DTT, 0.1 mM dNTPs, 6 U of *E.coli* DNA ligase, 6 U of DNA polymerase I, 1 U of *E.coli* RNase H, RNase-free water up to 50 µl, was incubated at 16 °C for 1.5 h followed by 70 °C for 10 min to terminate the reaction.

Digestion and ligation of ds-cDNA

Digestion and ligation reaction were performed on a O.T.P. Safety thermostat (PolyScience, USA). The 5 µl of digestion mixture containing 2.5 µl of ds-cDNA solution, 0.5 µl of 10×K Buffer, 5 U of *Mbo* I endonuclease, 1.5 µl of sterilized water, was mixed gently and incubated at 37 °C for 2 h followed by 75 °C for 15 min to terminate the reaction. Then digested ds-cDNA was ligated to source-specific adapters, respectively. The sequences of universal primers, gene-specific primers and "Y" shaped adapters were listed in Table S4†. The 20 µl of ligation reaction mixture containing 4 µl of the digested mixture, 2 µl of sterilized water, and 20 pmol of a source-specific adapter was incubated at 70 °C for 10 min, then 2 µl of 10 U/µl T4 DNA ligase and 2 µl of 10×T4 DNA ligase buffer were added before the incubation at 16 °C for 2 h. The ligase was inactivated at 75 °C for 15 min. Finally, equal volume of the barcode-ligated fragments from different sources were pooled for PCR amplification.

PCR and purification of PCR product

PCR amplification was performed on the PTC-225 Thermal Cycler PCR system (MJ Research, Inc., USA). Fifty microliters of PCR reaction containing 1.5 mM $MgCl_2$, 0.2 mM of each dNTP, 1.2 µM of each gene-specific primer and the universal primer, 1 µl of equally mixed ligation products as templates, and 1.25 U of *rTaq* DNA polymerase. The PCR program is as follows: denatured at 94 °C for 3 min, followed by 35 thermal cycles (94 °C for 30 s; 55 °C for 30 s; 72 °C for 45 s). After the cycle reaction, the product was incubated at 72 °C for 10 min and held at 16 °C. Amplification was performed on a PTC-225 Thermal Cycler PCR system (MJ Research, Inc., USA). Then the PCR products were purified by BioSpin PCR purification kit to remove surplus biotinylated primers.

As pyrosequencing was used for decoding the barcodes in PCR products, cross-contamination from PCR products may occur. To prevent any possible false positive results, it is necessary to perform PCR using filter-tips at a qualified PCR lab. In addition, a negative control should be carried out in parallel from the beginning of the assay (enzyme digestion step). Agarose gel-electrophoresis can be used to verify the negative control. If a clear pattern is obtained, pyrosequencing is then performed on samples.

Template preparation for pyrosequencing

The biotinylated PCR products were immobilized onto streptavidin-coated M280

Dynabeads, and DNA strands were separated by melting with 0.1 M NaOH. After washing with 1×annealing buffer (4 mM Tris-HCl, pH 7.5, 2 mM $MgCl_2$, 5 mM NaCl), the sequencing primer was added to the single-stranded DNA for annealing at 80 °C for 2 min.

Pyrosequencing reaction

Pyrosequencing was carried out by the reported method.[1-8] The reaction volume was 100 µl, containing 0.1 M tris-acetate (pH 7.7), 2 mM EDTA, 10 mM magnesium acetate, 0.1% BSA, 1 mM dithiothreitol, 2 µM adenosine 5'-phosphosulfate 0.4 mg/ml PVP, 0.4 mM D-luciferin, 200 mU/ml ATP sulfurylase, 3 µg/ml luciferase, 18 U/ml Sequenase 2.0 or 18 U/ml Exo^- Klenow Fragment, and 1.6 U/ml apyrase.

Pyrosequencing can also be performed by using commercialized PyroMark GOLD reagent kit (Qiagen, Germany) at PyroMark ID instrument (Qiagen, Germany).

Real-time PCR

Real-time PCR amplification was carried out with a DNA Engine Opticon 2 Thermal Cycler system (MJ Research, Inc., USA). Twenty-five µl of reaction mixture contains 12.5 µl of SYBR *Premix Ex* Taq™ (2×), 0.4 µM each primer, 1 µl of diluted (1/30) ss-cDNA (synthesized from 1.5 µg of total RNA), 9.5 µl of sterilized water. The program used was initial denaturation at 94 °C for 5 min followed by 40 cycles of 94 °C for 30 s, 55 °C for 30 s,72 °C for 40 s, plate read at 81 °C for 0.01 s, followed by a final extension of 7 min at 72 °C.

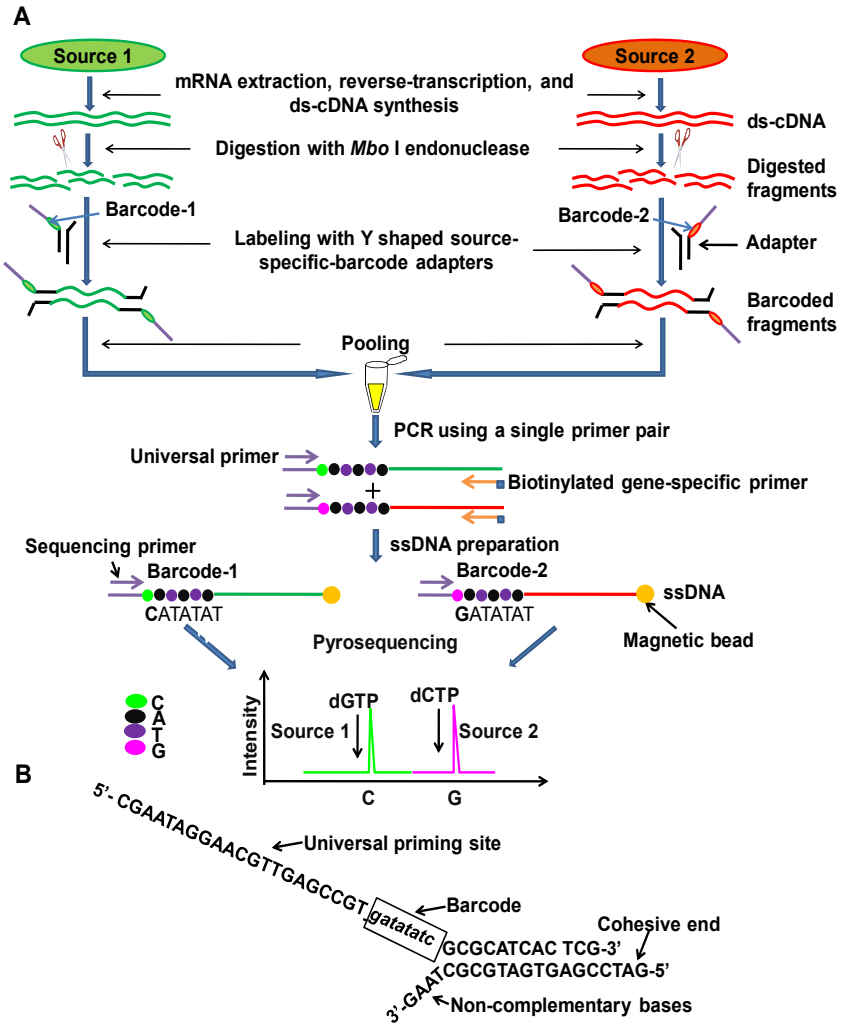

Fig. S1† Flowchart using to analyze relative expression levels of a given gene between two different sources (A) and the typical structure of "Y"-shaped source-specific-barcode adapter (B). **This figure illustrates the principle of BAMGA and the structure of "Y"-shaped source-specific-barcode adapter.**

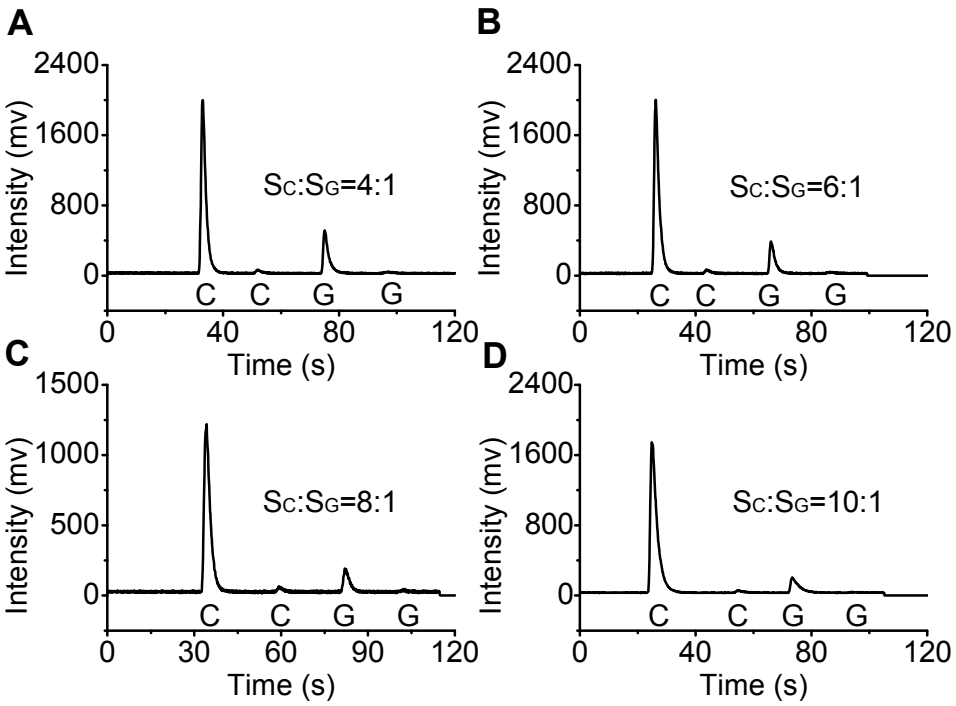

Fig. S2† Effect of coexisting DNA fragments on the accuracy of BAMGA. The theoretical ratios of the *β-ACTIN* gene between 2 artificially-prepared sources, which contains extra 0.42 μg of mouse-source double-stranded non-target cDNA, are 4:1 (A), 6:1 (B), 8:1 (C), and 10:1 (D), respectively. The expected ratios were directly marked in the pyrograms. The ratios were calculated with the peak intensities specific to barcodes encoded by adapter ligation reaction. **This figure indicates that non-target RNA coexisting in target RNA barely affect the accuracy of BAMGA.**

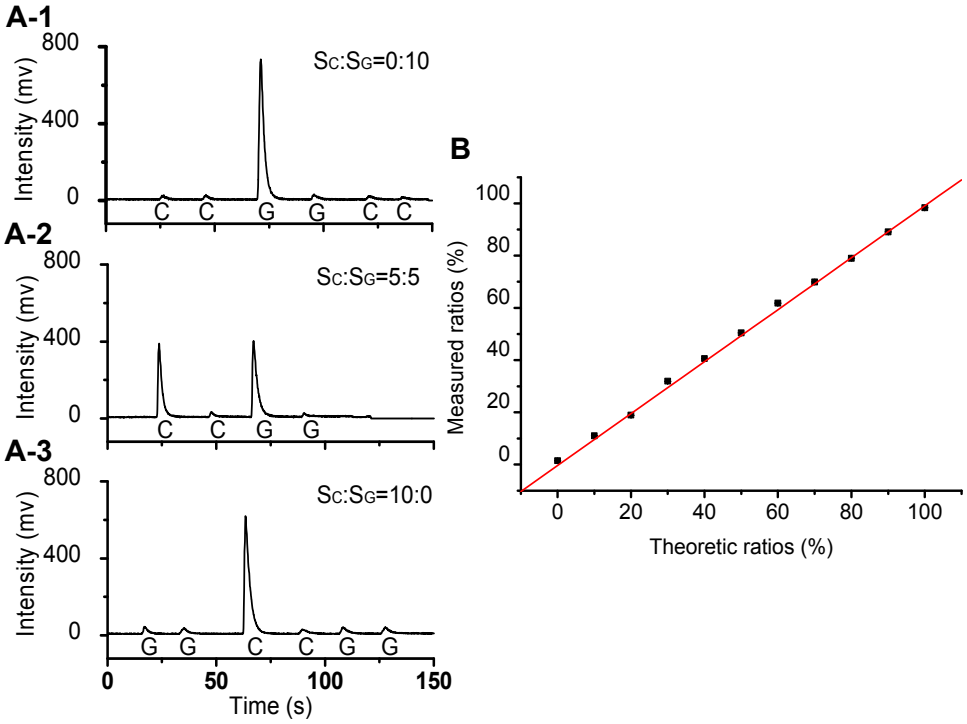

Fig. S3† Quantitative evaluation of the multiplex bioassay based on BAMGA. (A) Typical pyrograms of artificially-prepared sources by mixing ds-cDNAs ligated with barcode-1 and barcode-2 at various ratios (0:10, 5:5, and10:0). Each dNTP was dispensed twice for monitoring the background due to PPi impurity in dNTP solution. (B) Correlation between the expected ratios of the amounts of the *GAPDH* transcripts in two sources and the measured ratios by BAMGA. The ratios were calculated with the formula of "G/ (G + C) × 100" where G and C mean the peak intensity. **This figure indicates that BAMGA owns a good quantification performance.**

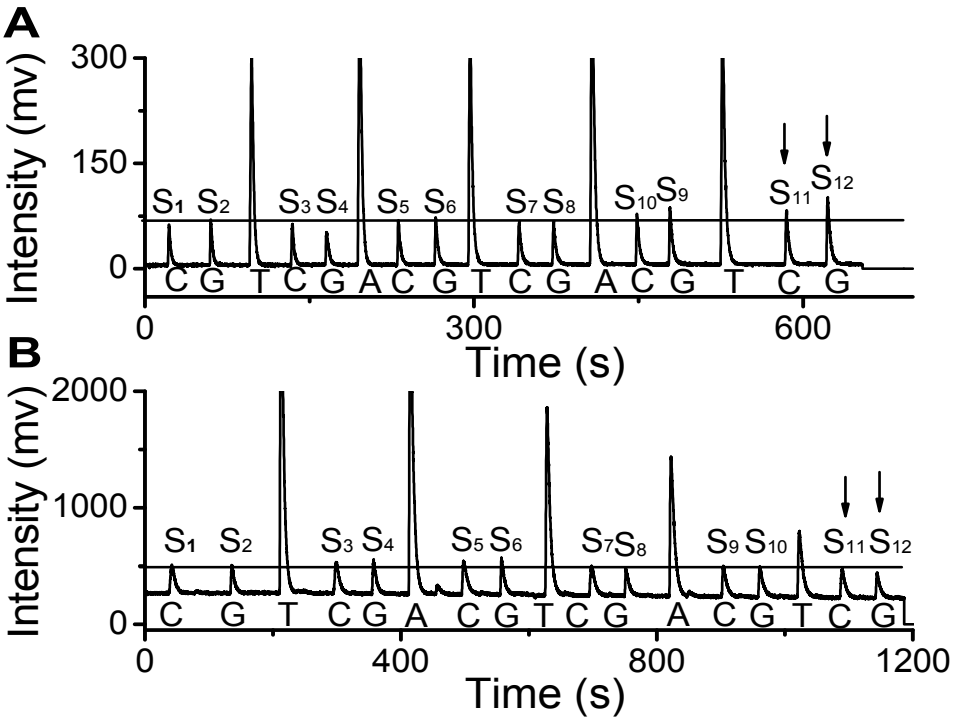

Fig. S4† Pyrograms for decoding barcodes by using dNTPs (A) and ddNTPs (B) in pyrosequencing reaction. The transcript concentration of the target gene (the *CDK4* gene) in each of 12 artificially prepared sources (from S1 to S12) is equal. **This figure indicates that pyrosequencing by ddNTP improves the quantification quality of BAMGA.**

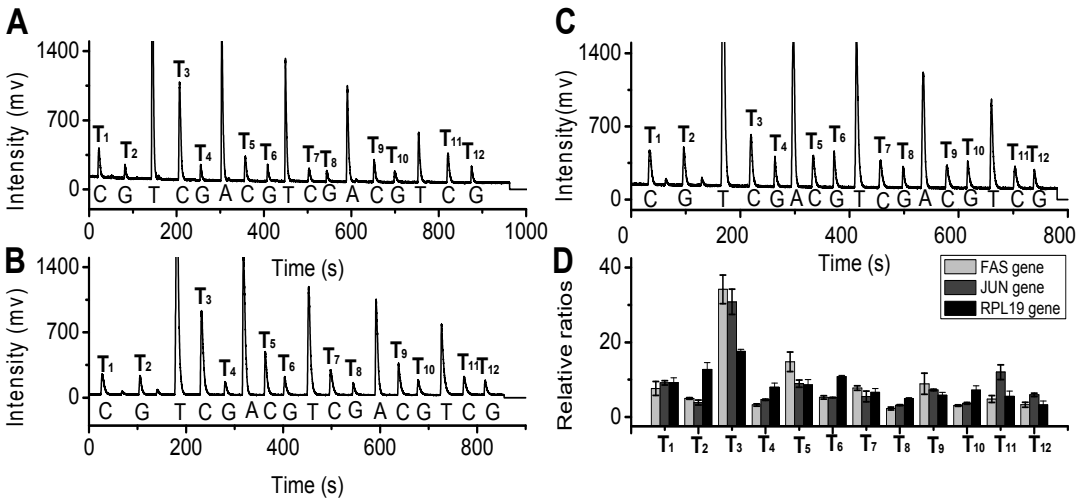

Fig. S5†. Results of relative gene expression levels of the *JUN* gene (A), the *RPL19* gene (B), and the *FAS* gene (C) among 12 different organs (small intestines, right brain, large intestine, cerebellum, spleen, lung, left brain, bladder, heart, stomach, ear, and kidney) of a mouse by multiplex bioassay based on BAMGA. "**T**" means the tissue in which a gene of interest exists. (D) Histograms of relative gene expression levels of the three genes among 12 different organs of a mouse by pyrosequencing. **This figure indicates the data of 12-plex BAMGA for detecting real biological samples.**

Table S1†. Results of gene expression levels of the *β-ACTIN* gene among 12 different organs of a mouse by real-time PCR.

Organs	CT value	Copies	Average copies
small intestines	23.127	138444	145652
	23.353	121952	
	22.693	176561	
right brain	23.191	133575	91163
	24.187	76384.5	
	24.515	63530.6	
large intestine	22.035	255466	253477
	22.127	242664	
	21.988	262302	
cerebellum	24.149	78026.6	67881
	24.936	50154.2	
	24.208	75463.1	
spleen	21.825	287418	269926
	22.135	241504	
	21.866	280856	
lung	23.958	86840.8	110280
	23.318	124333	
	23.387	119666	
left brain	23.980	85779.7	83260
	24.247	73840.8	
	23.891	90159	
bladder	24.335	70266.3	68633
	24.174	76934.2	
	24.656	58698.1	
heart	26.596	19766.4	14997
	27.633	11044.9	
	27.188	14180.2	
stomach	24.188	76308.7	61711
	24.565	61763.4	
	25.050	47060.9	
ear	27.273	13519.9	14930
	26.981	15926.5	
	27.047	15346.9	
kidney	24.364	69159.5	53973
	25.400	38664.5	
	24.802	54096.1	

Table S2†. Comparison of the relative expression levels of the *β-ACTIN* gene among 12 organ tissues of a mouse between BAMGA and real-time PCR.

Methods	Small intestines (T_1 %)	Right brain (T_2 %)	Large intestine (T_3 %)	Cerebellum (T_4 %)	Spleen (T_5 %)	Lung (T_6 %)	Left brain (T_7 %)	Bladder (T_8 %)	Heart (T_9 %)	Stomach (T_{10} %)	Ear (T_{11} %)	Kidney (T_{12} %)
Real-time PCR	12.3±1.8	7.6±2.4	21.6±3.2	5.8±1.7	22.8±1.1	9.5±2.6	7.1±1.1	5.8±0.9	1.2±0.2	5.1±0.6	1.3±0.1	4.5±1.1
BAMGA	13.7±1.8	8.4 ±1.0	19.2±3.5	6.1 ±0.4	21.0±1.1	7.6±0.3	6.3±0.7	5.1±0.8	2.4±0.3	4.4±0.7	2.9±1.3	3.0±1.2

Table S3†. Results of relative expression levels of the *β-ACTIN*, *JUN*, *RPL19*, and *FAS* gene among 12 organ tissues of a mouse by BAMGA.

Genes	Small intestines (T_1 %)	Right brain (T_2 %)	Large intestine (T_3 %)	Cerebellum (T_4 %)	Spleen (T_5 %)	Lung (T_6 %)	Left brain (T_7 %)	Bladder (T_8 %)	Heart (T_9 %)	Stomach (T_{10} %)	Ear (T_{11} %)	Kidney (T_{12} %)
β-ACTIN	13.7±1.8	8.4 ±1.0	19.2±3.5	6.1 ±0.4	21.0±1.1	7.6±0.3	6.3±0.7	5.1±0.8	2.4±0.3	4.4 ±0.7	2.9±1.3	3.0±1.2
JUN	9.1±0.6	3.9±0.7	30.8±3.4	4.6±0.2	8.9±1.0	5.2±0.1	5.5±1.4	3.1±0.1	7.2±0.2	3.6±0.3	12.0±2.0	5.9±0.4
RPL19	9.2±1.3	12.7±1.9	17.5±0.6	7.9±1.2	8.6±1.4	10.8±0.4	6.6±1.1	4.9±0.3	5.8±0.8	7.2±1.1	5.5±1.4	3.2±1.0
FAS	7.6±1.9	5.0±0.2	34.2±3.9	3.2±0.3	14.8±2.6	5.2±0.4	7.7±0.6	2.2±0.4	8.9±2.8	3.1±0.2	4.8±0.9	3.3±0.7

Table S4†. Primers for relative gene expression analysis by multiplex bioassay based on BAMGA.

Primers	Sequences (5'→3')
adaptor-1-G	CGAATAGGAACGTTGAGCCGTGATATATCGCGCATCACTCG
adaptor-1-C	CGAATAGGAACGTTGAGCCGTCATATATGGCGCATCACTCG
adaptor-2-G	CGAATAGGAACGTTGAGCCGTTGTATAACGCGCATCACTCG
adaptor-2-C	CGAATAGGAACGTTGAGCCGTTCTATAAGGCGCATCACTCG
adaptor-3-G	CGAATAGGAACGTTGAGCCGTTAGATATCGCGCATCACTCG
adaptor-3-C	CGAATAGGAACGTTGAGCCGTTACATATGGCGCATCACTCG
adaptor-4-G	CGAATAGGAACGTTGAGCCGTTATGTAACGCGCATCACTCG
adaptor-4-C	CGAATAGGAACGTTGAGCCGTTATCTAAGGCGCATCACTCG
adaptor-5-G	CGAATAGGAACGTTGAGCCGTTATAGATCGCGCATCACTCG
adaptor-5-C	CGAATAGGAACGTTGAGCCGTTATACATGGCGCATCACTCG
adaptor-6-G	CGAATAGGAACGTTGAGCCGTTATATGACGCGCATCACTCG
adapter-6-C	CGAATAGGAACGTTGAGCCGTTATATCAGGCGCATCACTCG
adapter-UP	GATCCGAGTGATGCGCTAAG
Up-1	CGAATAGGAACGTTGAGCCGT
bio-Up-1	bio-CGAATAGGAACGTTGAGCCGT
Up-2	CGAATAGGAACGTTGAGC
bio-Up-2	bio-CGAATAGGAACGTTGAGC
Mus-GAPDH-F	ATGTCGTGGAGTCTACTGGTGT
Mus-GAPDH-R	TGGAAGAGTGGGAGTTGCTG
Mus-CDK4-F	CGTTGGCTGTATCTTTGCAG
Mus-CDK4-R	AGGAGAGGTGGGGACTTGTT
Mus-β-ACTIN-F	CTGTCCCTGTATGCCTCTGGT
Mus-β-ACTIN-R	CTTTGATGTCACGCACGATTT
Mus-RPL19-F	GCCTGTGACTGTCCATTCCC
Mus-RPL19-R	CTTGTCTGCCTTCAGCTTGTG
Mus-FAS-F	GGACCTTGGAAAATCAACCC
Mus-FAS-R	GCAGAGAATAACAGAAGCAGT
Mus-JUN-F	GATTTGAAGCATTCTGAGTTTACC
Mus-JUN-R	ACACCACACCATCTTCTGGTG

The underlined bases represent the sequences of barcodes.

References

1. F. Mashayekhi and M. Ronaghi, *Anal Biochem*, 2007, **363**, 275-287.
2. A. M. Divne, H. Edlund and M. Allen, *Forensic Sci Int Genet*, **4**, 122-129.
3. B. Gharizadeh, J. Eriksson, N. Nourizad, T. Nordstrom and P. Nyren, *Anal Biochem*, 2004, **330**, 272-280.
4. B. Gharizadeh, T. Nordstrom, A. Ahmadian, M. Ronaghi and P. Nyren, *Anal Biochem*, 2002, **301**, 82-90.
5. T. Nordstrom, K. Nourizad, M. Ronaghi and P. Nyren, *Anal Biochem*, 2000, **282**, 186-193.
6. G. Zhou, M. Kamahori, K. Okano, K. Harada and H. Kambara, *Electrophoresis*, 2001, **22**, 3497-3504.
7. B. Gharizadeh, A. Ohlin, P. Molling, A. Backman, B. Amini, P. Olcen and P. Nyren, *Mol Cell Probes*, 2003, **17**, 203-210.
8. J. Eriksson, B. Gharizadeh, T. Nordstrom and P. Nyren, *Electrophoresis*, 2004, **25**, 20-27.