Serial pyrosequencing for quantitative DNA methylation analysis

Jörg Tost, Hafida El abdalaoui, and Ivo Glynne Gut

Centre National de Génotypage, Evry, France

In mammals, DNA methylation is a genetically programmed DNA modification and occurs almost exclusively at the 5-position of the pyrimidine ring of cytosines. Cytosine methylation is essential for mammalian embryogenesis and development, implicated in X inactivation, and influences gene expression of the corresponding genes, giving rise to distinct expressions patterns in different cell and tissue types (1). Relatively CpG-rich clusters of approximately 1–2 kb length, the CpG islands, are found in the promoter regions and first exons of many genes and are mostly not methylated. Aberrant patterns of DNA methylation have been identified in the pathogenesis of various human diseases. DNA methylation exhibits tumor-specific patterns, and epigenetic changes occur at least as often as genetic mutations in carcinogenesis (2). Epimutations appear early in the progression of cancer and often precede malignancy. Due to the importance of epigenetic changes in development and disease, a variety of techniques for the study of DNA methylation has been developed in recent years (2,3). The choice of the method mainly depends on the desired application. No single method has emerged as a gold standard unifying quantitative accuracy, high sensitivity, potential for whole genome analysis, and precise investigations of individual CpG positions.

Pyrosequencing is a real-time DNA sequencing-by-synthesis method that relies on an enzymatic cascade that quantitatively converts the pyrophosphate released upon nucleotide incorporation into a luminometric signal (4). It is ideally suited for DNA methylation analysis after bisulfite treatment, as it combines the capability for direct quantitative sequencing with reproducibility, speed, and ease-of-use (5–10). Limitations are the length of the sequence read, and thereby the number of CpGs that can be analyzed in one sequencing reaction, and the relatively high cost that is equally shared between the PCR amplification step, the pyrosequencing reaction, and, to a lesser extent, the purification step. Conventional multiplex detection using several sequencing primers is not feasible for DNA methylation analysis due to the low sequence complexity of bisulfite-treated genomic DNA. Here we report significant improvements to the current protocol for DNA methylation analysis by pyrosequencing that reduce cost, labor, and analysis time, as well as saving precious DNA samples from clinical specimens.

Genomic DNA samples from the Centre d'Etude du Polymorphisme Humain (CEPH)/Utah collection were obtained from the DNA bank of the Centre National de Génotypage (Evry, France) and bisulfite-treated as

Table 1. Sequences of Primers Used for Amplification, Pyrosequencing Reactions, and the Dispensation Order for the Respective Analyses

Gene	PCR Primer Forward	PCR Primer Reverse	Pyrosequencing Primer	Dispensation Order	CpGs
HLA-G (218 bp)	5'-GTGTTT-TTTAGGTAG-GTTTTTTTTT-3'	5'-biotin-AATTCT-CACACCCTC-CAATAAATAAT-3'	5'-TTTTTTGTTTAGT-TATATT-3'	5'-CGTCAGTTATATGTCAGTGAGATGAGTCAG-TAGTGTCTGTGTCAGTAGAGTCAT-3'	8–13
			5'-GCGAGGTA-ATTTTTGTTAT-3'	5'-ATCAGTAGTATATGTATATCAGTCTGAGAGT-CAGTCAGTCAGATTAGTCAGT-3'	16–22
TABPB (119 bp)	5'-TGGTGTTT-TATTTTGAA-TTTTTTGTT-3'	5'-biotin-TACCAT-TATAAAAAAACCT-TCTCCC-3'	5'-TTTTGAATTTTTT-GTTGTAT-3'	5'-GATCAGGTAGTATCAGTCTG-3'	14–16
			5'-AGATTGAGGTG-TATGTTTA-3'	5'-ATCAGTGTCAGTCTG-3'	17–19
IGF2 DMR2 (255 bp)	5'-GGGAAA-GGGGTTTAG-GATTTTTAT-3'	5'-biotin-ATATT-TACTCCCCCT-TCAACCTC-3'	5'-GATTTTTATYG-GAAGTA-3'	5'-GTCTGTCGAGGTCGATGATCAGTTCGA-3'	2–6
			5'-GGGTYGATAY-GTTTTTTT-3'	5'-GTCGATAGTCGGTAGTATAGTGATCGTTAGAG-TAGTATAGTAGTCGATAGTATTGATCGATGTGATC-GTAGTCGTATCGT-3'	6–13
			5'-AGTATTTTTTAAT-GATGCTA-3'	5'-ACTCGTGCTCGACTCGTATACTCGTGCTCG-3'	11–15
			5'-TTTTATAG-TATAGAGAGAGT-3'	5'-ATCGTGTAGTGATCGTATGATTATGTTCGA-3'	16–19

CpGs are numbered in the order of appearance from the 5' end of an amplification product. *HLA-G,* human leukocyte antigen-G; *TABPB,* human transporter associated with antigen processing binding protein; *IGF2,* human insulin-like growth factor 2; *DMR2,* differentially methylated region 2.

described in detail elsewhere (8). Amplified targets comprised (*i*) a fragment of the CpG island spanning the transcription start of the human leukocyte antigen-G (*HLA-G*) gene, (*ii*) a fragment of the CpG island spanning the transcription start of the human transporter associated with antigen processing binding protein (*TAPBP*), and (*iii*) a portion of the differentially methylated region 2 (DMR2) of the imprinted human insulin-like growth factor 2 (*IGF2*). *HLA-G* was amplified using 20 ng bisulfite-treated human genomic DNA and 5 pmol forward and reverse primer, the latter being a biotinylated primer (Table 1). The reaction conditions were 1× HotStarTaq® buffer supplemented with 1.6 mM MgCl$_2$, 200 µM dNTPs, and 2.0 U HotStarTaq polymerase (Qiagen, Hilden, Germany) in a 25-µL volume. For *HLA-G*, the PCR program consisted of a denaturing step of 15 min at 95°C, followed by 50 cycles of 95°C for 30 s, 62.2°C for 30 s, and 72°C for 20 s, with a final extension of 72°C for 5 min. *TAPBP* was amplified with the appropriate forward and reverse primers (again, the latter primer is biotinylated). The reaction conditions were 60 mM Tris-SO$_4$, pH 8.9, 18 mM (NH$_4$)$_2$SO$_4$, 1 mM MgCl$_2$, 200 µM dNTPs, and 2.0 U Platinum *Taq* polymerase in a 25-µL volume. The PCR program consisted of a denaturing step of 95°C for 4 min, followed by 50 cycles of 95°C for 30 s, 57.5°C for 30 s, and 72°C for 20 s, with a final extension of 72°C for 5 min. The amplification of the *IGF2* target region was carried out as previously described (8). PCR products were incubated for 10 min at room temperature with 51 µL binding buffer (10 mM Tris, 2 M NaCl, 1 mM EDTA, 0.1% Tween® 20, pH 7.6, adjusted with 1 M HCl) and 4 µL streptavidin-coated Sepharose™ beads (GE Healthcare, Uppsala, Sweden). The binding mixture was aspirated using the Pyrosequencing® Vacuum Prep Workstation (Biotage, Uppsala, Sweden) (11), and the template was successively washed with 70% ethanol, rendered single-stranded with 0.2 M NaOH, and neutralized with washing buffer (10 mM Tris, pH 7.6, adjusted with 4 M acetic acid). Beads
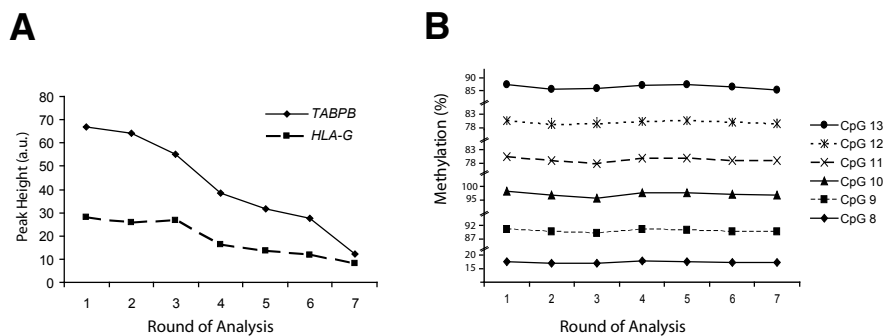

Figure 1. Signal intensity and quantification of methylation during serial pyrosequencing. (A) Course of signal intensity for a sequencing primer in the human transporter associated with antigen processing binding protein (*TAPBP*) promoter and the first exon of human leukocyte antigen-G (*HLA-G*) gene during repeated stripping and reannealing of a sequencing primer. Signal intensity is given in arbitrary units (a.u.), taking into account seven peaks that correspond to single nucleotide dispensations in a de novo synthesized DNA strand. (B) Quantitative DNA methylation analysis of six CpG positions in the first exon of *HLA-G* during seven rounds of repeated stripping of the sequencing primer. Quantification is robust down to a single peak height of 10 a.u., which was used as a cut-off value.
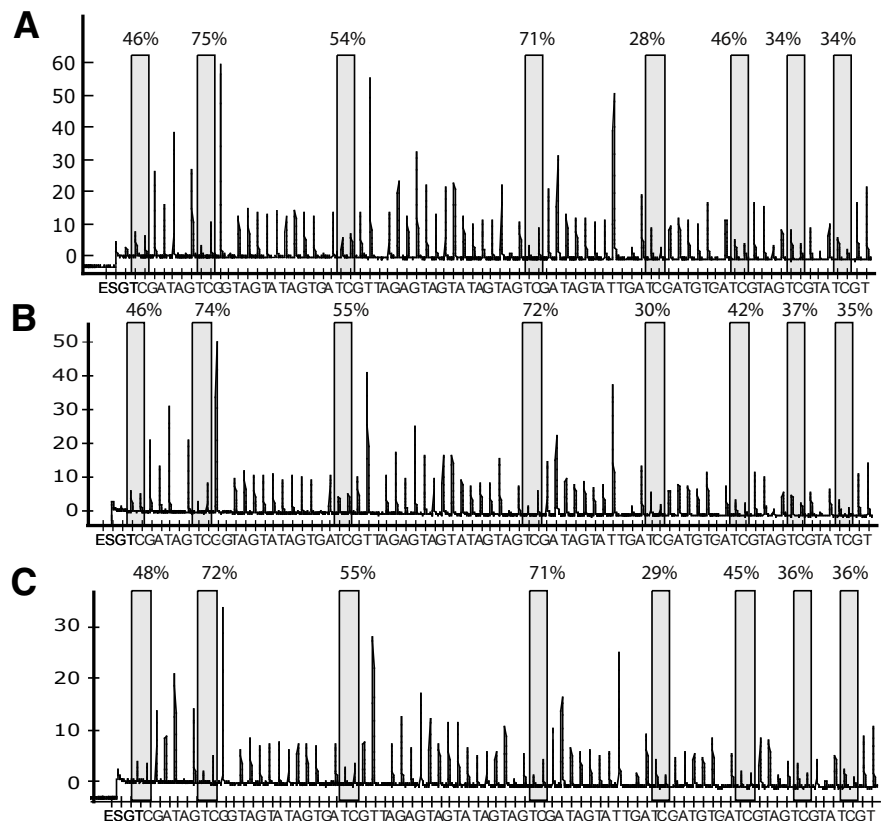

Figure 2. Quantitative DNA methylation analysis of eight CpGs in the human insulin-like growth factor 2 (*IGF2*) differentially methylated region 2 (*DMR2*). A sample corresponding to an individual from the Centre d'Etude du Polymorphisme Humain (CEPH) DNA collection was amplified in quadruplicate and analyzed with the four sequencing primers. Shown are three results for sequencing primer 5′-GGGTYGATAYGTTTTTTTT-3′ (A) in the first analysis on the first amplification product, (B) in the second analysis (after one round of stripping) on a second PCR product, and (C) in the fourth analysis on the fourth amplification product, respectively. Although signal intensity decreases, quantitative DNA methylation values are constant, demonstrating the robustness of the presented method. Quantitative analysis of CpG positions in regions covered by two separate sequencing primers showed good concordance (maximum deviation of 5%) between the two measurements independent of the round of analysis.

were released into 40 µL annealing buffer (20 mM Tris, 2 mM magnesium-acetate, pH 7.6, and adjusted with 4 M acetic acid) containing 15 pmol of the respective sequencing primer (Table 1). Primers were annealed to the target by incubation at 80°C for 2 min. Quantitative DNA methylation analysis was carried out on a PSQ 96MA system with the Pyro Gold sequence analysis (SQA) reagent kit (Biotage), and results were analyzed using the Q-CpG software (V.0.1.7; Biotage). Stripping of the template strand for subsequent annealing of a new sequencing primer was carried out by adding 20 µL binding buffer to the completed sequencing reaction and resuspending the Sepharose beads. The binding mixture was then purified without further incubation, and the biotinylated template strand was rendered single-stranded again using the purification protocol described above. This process completely removes all DNA strands that have been de novo synthesized during the last sequencing run, as well as the remaining sequencing primers. To test for the maximum number of sequencing runs possible on the same single-stranded PCR products, the *HLA-G* fragment was analyzed by alternately switching between the two different sequencing primers. Pyrosequencing can be repeated up to seven times without any loss of accuracy (Figure 1). PCRs were performed in triplicate, and peak height measurements were averaged between those technical replicates. Although there is a slight loss in intensity after each round of stripping in peak height, probably due to incomplete recuperation of the Sepharose-coated beads, the quantification of the methylation is reproducible between the pyrosequencing analyses (Figure 1B). The quantitative methylation values vary over a range of up to 5% (average 2.7%) over the course of the experiment, which is also the range of the standard deviation obtained by independent PCR amplification of the same sample. No tendency toward the major or minor allele was observed over the different analyses (Figure 2). In control experiments, when the new sequencing primer is omitted in the

following pyrosequencing experiment, no signals corresponding to the incorporation of nucleotides are detected.

An added advantage of this modification of the established pyrosequencing protocol is the reduction of the required volumes of enzyme and substrates of the Pyrosequencing SQA kit. Since the cartridge can be reused without washing between consecutive runs, the dead volume of the cartridge (60 µL) can be subtracted from the required volumes. The described method thereby significantly reduces reagent costs and labor time and accelerates the workflow, as time-consuming PCRs have to be performed only once. This also reduces the risk of contamination, an inherent problem of PCR amplification with a large number of amplification cycles. A major advantage of the protocol is the conservation of precious DNA resources that can be used for the analysis of more genes rather than being used up for scanning fewer fragments completely. The use of three to four sequencing primers in most cases permits the sequencing of the entire PCR product, as these are limited in their size to avoid formation of secondary structures that lead to an increase of background signals. The described method is particularly useful for the identification of methylation variable positions that might be used as diagnostic and/or prognostic epigenetic markers implicated in human disease.

ACKNOWLEDGMENTS

COMPETING INTERESTS STATEMENT

The authors declare no competing interests.

REFERENCES

1. **Bird, A.** 2002. DNA methylation patterns and epigenetic memory. Genes Dev. *16*:6-21.
2. **Laird, P.W.** 2003. The power and the promise of DNA methylation markers. Nat. Rev. Cancer *3*:253-266.
3. **Cottrell, S.E.** 2004. Molecular diagnostic applications of DNA methylation technology. Clin. Biochem. *37*:595-604.
4. **Ronaghi, M., M. Uhlén, and P. Nyren.** 1998. A sequencing method based on real-time pyrophosphate. Science *281*:363-365.
5. **Uhlmann, K., A. Brinckmann, M.R. Toliat, H. Ritter, and P. Nürnberg.** 2002. Evaluation of a potential epigenetic biomarker by quantitative methyl-single nucleotide polymorphism analysis. Electrophoresis *23*:4072-4079.
6. **Tost, J., J. Dunker, and I.G. Gut.** 2003. Analysis and quantification of multiple methylation variable positions in CpG islands by Pyrosequencing. BioTechniques *35*:152-156.
7. **Colella, S., L. Shen, K.A. Baggerly, J.P. Issa, and R. Krahe.** 2003. Sensitive and quantitative universal Pyrosequencing methylation analysis of CpG sites. BioTechniques *35*:146-150.
8. **Dupont, J.M., J. Tost, H. Jammes, and I.G. Gut.** 2004. De novo quantitative bisulfite sequencing using the pyrosequencing technology. Anal. Biochem. *333*:119-127.
9. **Yang, A.S., M.R. Estecio, K. Doshi, Y. Kondo, E.H. Tajara, and J.P. Issa.** 2004. A simple method for estimating global DNA methylation using bisulfite PCR of repetitive DNA elements. Nucleic Acids Res. *32*:e38.
10. **Shaw, R.J., T. Liloglou, S.N. Rogers, J.S. Brown, E.D. Vaughan, D. Lowe, J.K. Field, and J.M. Risk.** 2006. Promoter methylation of P16, RARbeta, E-cadherin, cyclin A1 and cytoglobin in oral cancer: quantitative evaluation using pyrosequencing. Br. J. Cancer *94*:561-568.
11. **Dunker, J., U. Larsson, D. Petersson, J. Forsell, A.L. Schiller, A. Alderborn, and L.M. Berg.** 2003. Parallel DNA template preparation using a vacuum filtration sample transfer device. BioTechniques *34*:862-868.

Address correspondence to Jörg Tost, Centre National de Génotypage, Bâtiment G2, 2 rue Gaston Crémieux, CP 5721, 91057 Evry Cedex, France. e-mail: tost@cng.fr

To purchase reprints

of this article, contact

Reprints@BioTechniques.com