Direct comparison of two different strategies to sequence bisulfite-treated genomic DNA for DNA methylation analysis in genetic diseases and cancer samples.

Mario F.Fraga[1], Alain Rico[2], Applera France, Britta Finkelnburg[3] and Manel Esteller[1]

[1]Cancer Epigenetics Laboratory; Spanish National Cancer Centre, Madrid (CNIO) Spain. [2]Applera France, Applied Biosystems.

[3]Applera Deutschland GmbH, Applied Biosystems.

***Corresponding authors:** mffraga@cnio.es, britta.finkelnburg@eur.appliedbiosystems.com.*

ABSTRACT

Enzymatic methylation on the cytosine of genomic DNA (gDNA) has been shown to correlate with gene expression. Specifically, methylation of cytosines (C) at CpG motifs, usually in the promoter regions of genes in CpG-dense regions known as CpG islands, will shut down expression of the gene in complex biological processes.Biological relevance is given for all developmental processes including aging and tumor growth and methylation pattern may be different between cell types.

A well known method to study methylation patterns is to treat gDNA by bisulfite to distinguish methylated cytosine (5mC) from unmethylated C, which is deaminated to uracil (U) and replaced by thymine (T) in subsequent amplification. 5mC still remains as C. Subsequent amplification can focus on selective amplification of methylation patterns in CpG islands (methyation specific PCR, MSP) or on amplification of bisulfite treated (converted) gDNA (Bisulfite treatment specific PCR, BSP). Selection of PCR focus is done by primer design. After PCR, sequencing can clarify the methylation pattern.

Due to very imbalanced nucleotide composition after bisulfite treatment bias in amplification of bisulfite treated gDNA as well as in sequencing products may lead to experimental problems. PCR improvements during amplification of bisulfite-converted gDNA by the use of universal tailed primers, combined with changes in thermal cycling and reactant concentrations are described earlier [1,2]. These PCR products can be sequenced by using the universal primer tails. This method will be compared to the cloning of BIS PCR products for sequencing the bacterial clones.

Summary of Key findings

1. Methyl Primer Express® Software provides a successfull design for robust amplification of short and long amplicons.
2. Universal tailed Primers improve successful generation of results in general.
3. Strategy is depending on preferred outcome [more or less vs semi-quantitative] and sequence composition itself.
4. For Mixed-Base CpG Typing high specificity of PCR is a necessity.

Introduction

Due to general difficulties for MSP/BSP PCR, special recommendations for different approaches may be useful to maximize sucessful generation of data.

Sample tracking is depending on cell type and biological status of isolated cells. It could not be expected that all DNAs isolated will be in the same methylation state. If amplified, a population of molecules with different methylation pattern is very likely.

Bisulfite treatment is a very critical step. An incomplete conversion will lead to an artificial increase of different methylation patterns.

Depending on methylation state in general, a very imbalanced nucleotide composition after bisulfite treatment can be expected. This may bias amplification of all of bisulfite treated gDNA with different methylation patterns.

If a very C-rich sequence will be converted by bisulfite, long T-stretches will be the result. During amplification Taq Polymerases elongating the primers will slipp over the original template molecules in such areas. Results are populations of product molecules with different length of these T-stretches. Finally this may lead to problems during sequencing.

Amplicon length:

In general, shorter amplicons will be amplified more robust, bias is not that issue, errors are less frequent. This is an optimal cenario for sequencing PCR products. But with this strategy, it will only be possible to sequence the whole population of amplified molecules. All factors increasing the plurality of the population of molecules (including slippage errors) will directly influence the chance and the results of sequencing.

For covering large genomic regions, longer amplicons will be preferred for better coverage/less reactions. For this case, the likelyhood for more inefficient amplification and polymerase errors in A/T-rich regions can be tolerated, if down-stream analysis will be done by cloning. Separation of the population of amplified molecules will allow the sequencing of the individual fragments.

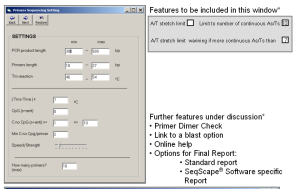
For both sequencing strategies application specific recommendations from primer design to analysis of data may be useful to maximize successful generation of results.

Materials and Methods

Amplicon Design: PCR primer design was done with MethylPrimer Express® Software following specific settings. **Bisulfite treatment**: the conventionally used method discribed in [3] was compared with the Bisulfite Conversion Kit from Applied Biosystems "currently in development". **PCR of bisulfite treated gDNA**: a setup with EcoStart5 [Ecogene, Spain] was compared to a setup with AmpliTaq Gold [Applied Biosystems, USA] using the GeneAmp 2720 Thermal Cycler.**Cloning and DNA preparation**: method discribed in [3]. **Sequencing of DNA from clones**: BigDye™ Terminator v3.1 chemistry was used for sequencing reactions performed in a GeneAmp 9700 Thermal Cycler. Samples were cleaned up by ethanol precipitation. **Cleanup of PCR products and direct sequencing with universal primers**: after BSP PCR primers and dNTPs were removed by Exo/SAP-IT® (USB Corporation,USA) treatment. BigDye™ Terminator v1.1 chemistry was used for sequencing reactions performed in a GeneAmp 9700 Thermal Cycler. Samples were cleaned up by by ethanol precipitation. **Separation of sequences**: Sequences optained from clones were separated on Applied Biosystems 3700 DNA Analyzer running with POP-6™ polymer and a 50 cm array. Primary analysis of data was done by Sequencing Analysis Software v3.7. Sequences optained from direct sequencing with universal primers were separated on Applied Biosystems 3130xl Genetic Analyzer running on POP-7™ polymer and a 36 cm array. Primary analysis of data was done by Sequencing Analysis Software v5.2 using the KB basecaller. Sequencing Scanner Software v1.0 was used for reviewing data. **Comparison of data**: SeqScape® Software v2.5 was used to get the final results focussing on CpG typing possibilities with this software.

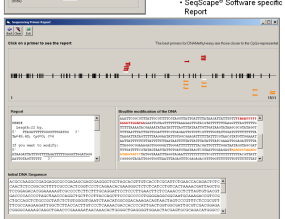
Amplicon Design with Methyl Primer Express® Software v 1.0

BSP for Cloning

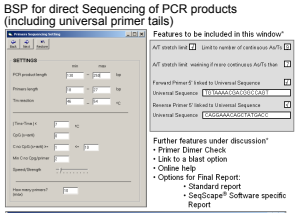


Features to be included in this window"

Further features under discussion"
- Primer Dimer Check
- Link to a blast option
- Online help
- Options for Final Report:
  - Standard report
  - SeqScape® Software specific Report

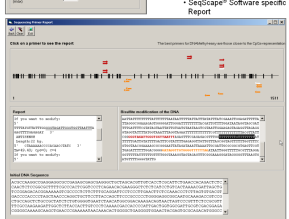


BSP for direct Sequencing of PCR products (including universal primer tails)



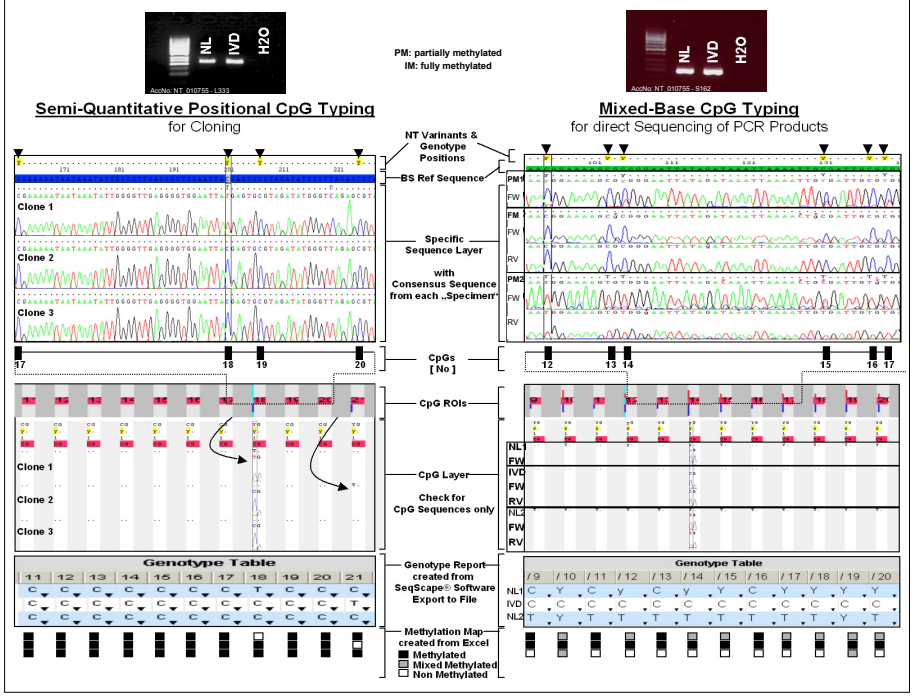
Features to be included in this window"

Further features under discussion"
- Primer Dimer Check
- Link to a blast option
- Online help
- Options for Final Report:
  - Standard report
  - SeqScape® Software specific Report

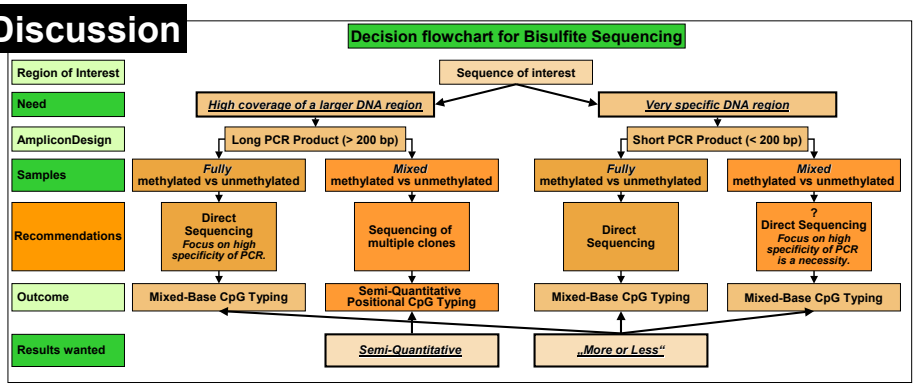


"Recommendations for design of BIS PCR for cloning and direct PCR sequencing with tailed primers are planned to be incorporated into a future version of the software by Applied Biosystems.

Results - Examples for different Strategies



Discussion



References and Trademarks

[1] Victoria L. Boyd and Gerald Zon. Analytical Biochemistry 326 (2004) 278–280. Bisulfite conversion of genomic DNA for methylation analysis: protocol simplification with higher recovery applicable to limited samples and increased throughput.

[2] Victoria Boyd, Diane Bond, Gerald Zon, Olga Petrauskene, and Cheryl Heiner. Poster ASHG 2004. Improved Protocols for the analysis of 5-methylcytosine in genomic DNA with bisulfite sequencing.

[3] Paz, M. F., Wei, S., Cigudosa, J. C., Rodriguez-Perales, S., Peinado, M. A., Huang, T. H. & Esteller, M. (2003) *Hum. Mol. Genet.* **12,** 2209–2219.

Abstract: 945/T