# Trends in Bioinformatics

ISSN 1994-7941

science alert

ANSInet
an open access publisher
http://ansinet.com

OPEN ACCESS Trends in Bioinformatics

ISSN 1994-7941
DOI: 10.3923/tb.2019.1.6

# Research Article
# Bisulfite Conversion DNA Visualizer for Designing DNA Methylation Primers

[1]Hernán G. Hernández, [2]Misael Leonardo López-Cepeda, [2]Gustavo A. Contreras-García, [2]Carolina Vargas-Porras and [3]Humberto Arboleda

[1]Department of Dentistry, Universidad Santo Tomás, 680002 Bucaramanga, Colombia
[2]Faculty of Health, Universidad Industrial de Santander, 680002 Bucaramanga, Colombia
[3]Genetics Institute, Universidad Nacional de Colombia, 111321 Bogotá, Colombia

## Abstract

**Background and Objective:** An important guideline for preventing amplification bias in Polymerase chain reaction (PCR) based DNA methylation techniques was recently proposed. However, there is no software that facilitates visualizing and applying such recommendations in PCR primers design. The aim of this work was to provide a web tool to facilitate the visualization of critical sequence elements for the design of primers applying the most recent recommendations for PCR-based methylation techniques. **Materials and Methods:** BiCVisualizer, a script for the visualization of critical sequence elements was designed using Java/html. The amplification specificity of a set of 10 primer pairs designed with the visual assistance of this web-tool was evaluated *in silico* PCR using BiSearch software. **Results:** The *in silico* PCR test showed a mean of 353 matches per primer and strand (range 84-934) using BiSearch. BiSearch did not find non-specific amplification for any primer pair designed using BiCVisualizer. **Conclusion:** This script was designed to assist the effective design of primers for methylation-sensitive high-resolution melting, the utility of this tool was supported by *in silico* evaluation. Users requiring visual construct primers can now spend less time in primer design using BiCVisualizer tool.

Key words: Java/html, DNA methylation, software, web tool, polymerase chain reaction

Citation: Hernán G. Hernández, Misael Leonardo López-Cepeda, Gustavo A. Contreras-García, Carolina Vargas-Porras and Humberto Arboleda, 2019. Bisulfite conversion DNA visualizer for designing DNA methylation primers. Trends Bioinform., 12: 1-6.

Corresponding Author: Hernán G. Hernández, Department of Dentistry, Universidad Santo Tomás, 680002 Bucaramanga, Colombia Tel: +573005683733



Competing Interest: The authors have declared that no competing interest exists.

Data Availability: All relevant data are within the paper and its supporting information files.

## INTRODUCTION

DNA methylation is considered a main epigenetic mark and has been associated with the pathogenesis of many diseases[1-3]. As bisulfite treatment changes DNA base composition of non-methylated Cytosines (Cs), it is used to study the methylation status of DNA sequences[4,5].

Methylation-Sensitive High-Resolution Melting (MS-HRM) is a bisulfite-based technique that allows for the detection of differentially Methylated Regions (DMRs) based on the assessment of DNA melting temperature. Melting temperature can be measured as sharp decreases in the fluorescence signal of an intercalating dye in real-time thermocyclers[6-8]. The MS-HRM is widely used for the detection of DMRs due to its high sensitivity, scalability, speed and low cost[9]. The MS-HRM protocols for the detection of DMRs have shown a measured limit of detection of 1% or even lower[6]. Therefore, there is a growing interest in the feasibility of using MS-HRM in cancer research and clinical diagnosis[6,10]. However, bisulfite-based Polymerase Chain Reaction (PCR) techniques including MS-HRM are prone to PCR bias, namely, the preferred amplification of one allele over another without biological meaning usually favoring non-methylated alleles[11]. To avoid such bias, Wojdacz *et al.*[11] made some recommendations for MS-HRM primers design. According to these recommendations some important critical sequence elements should be considered at the same time; however, current software does not permit the visualization of such elements, thus making the primer design for M S-HRM a time-consuming task[5].

Therefore, the objective of the present work was to develop a visualization tool to allow for the easy and quick selection of key-sequence elements in a user-friendly manner for applying the current recommendations of primer designing in PCR-based DNA methylation techniques. To obtain a functional tool that allows for the visualization of the mentioned key elements, the steps required were: Identification of the key elements in the DNA sequence, development of a java-html script to find and highlight the elements in an easy color code, systematically testing a set of primers designed with the developed tool through an *in silico* analysis and release of the software as a free available web-tool.

## MATERIALS AND METHODS

**Visualizer tool implementation:** The first step to develop the web tool was the identification of the key elements required in primer design according the Wojdacz's recommendations[11]. A shortened adaptation of the full set of original recommendations[12] was displayed in Table 1.

In the designed visualizer software, it considered the linguistic complexity reduction produced by 'bisulfite conversion' with disproportional decrease in 'Cs' content, in the original DNA strand or 'As' in the complementary strand. After bisulfite treatment, the methylated CpG dinucleotides represent the only reservoir where Cytosines remain Cytosines; therefore, most Cytosines had been depleted in the DNA sequence[13]. Consequently, a function for highlighting the high complexity sequences was included in BiCVisualizer. In sum, the following sequence elements were detected and highlighted to be visualized in BiCVisualizer: "CpG dinucleotides", "Thymines from a bisulfite-converted sequence", "mononucleotide repeats" and "linguistically complex tetranucleotide sequences".

**BiCVisualizer Java-html tool and on-line use:** BiCVisualizer tool for primer design was written for Java-html version 1.2. For public access the tool was made freely available online at http://dbic-visualizer.ml. For using BiCVisualizer, the input should be introduced in plain text or FASTA format with no characters other than A, G, C and T in DNA sequence.

The BiCVisualizer input sequence can be accepted in lowercase, uppercase, mixed lowercase/uppercase and with spaces or newline characters. In addition, practical user instructions are available at the BiCVisualizer's web site. Script designing and improving took the equivalent to 6 months of intensive work. Designing the testing primers took less than 30 min per primer pair.

The BiCVisualizer output is a bisulfate converted and highlighted version of the input sequence that is susceptible of being copied or modified in a text editor. The output sequence was intended to support users in designing primers according to the current recommendations (Fig. 1). The output sequence was designed for displaying a straightforward color code, assisting users to distinguish the following critical sequence elements:

Table 1: Summary of current Wojdacz's primer design recommendations and rationale

| Recommendations | Rationale |
|---|---|
| CpGs included should in the 5' end (faraway of 3') | Intended to control allele amplification bias |
| Include one CpG in the sequence of each primer | If bias continues, two CpGs could be included |
| "Inclusion of one or more Ts originating from a non-CpG C, at or near the 3' end" | Intended to control allele amplification bias |
| "The selected primers should be further evaluated in regard to standard parameters for primer design recommended" recommended TM ~60°C | Parameters of any primer should be evaluated before synthesis |

Source: Wojdacz *et al.*[11]

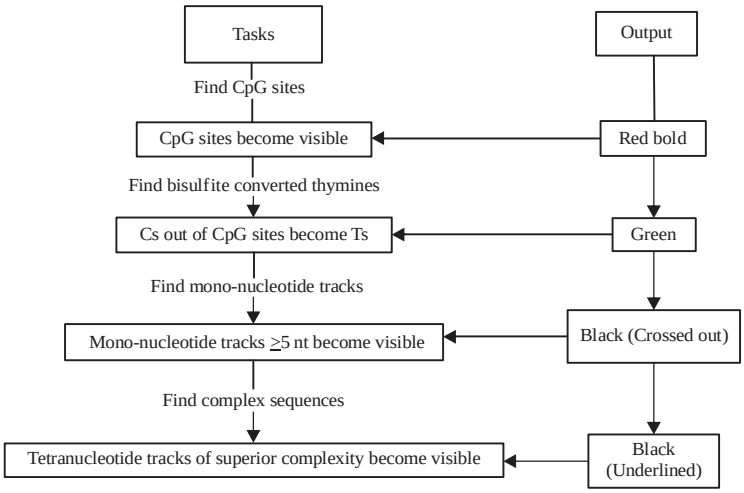


Fig. 1: Overall functions of BiCVisualizer

The diagram displayed the tool processes arrangement for the color/style code construction to make visible critical sequence elements for the design of the primers

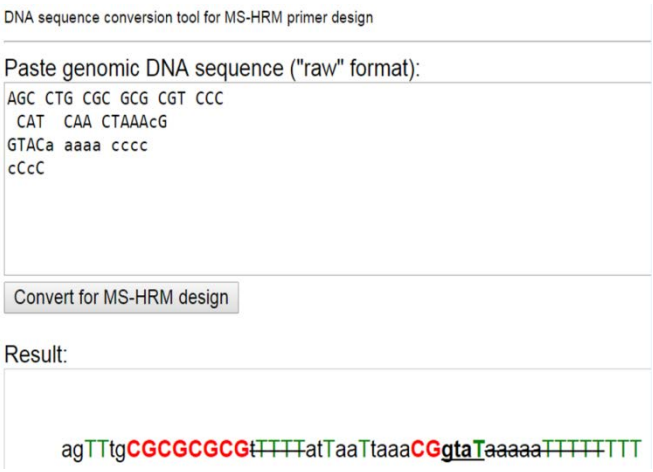


Fig. 2: Screen-shot of BiCVisualizer DNA visualizer for MS-HRM primer design web site interface

Upper square allows pasting sequence of interest in FASTA or raw format. Capital case letters, as well as lower case letters are allowed in the input, even if the letters are scattered or in separate lines. The bottom box shows the output bisulfite converted sequence. CpG sites are shown without conversion as if they were methylated in red boldcapital letters; the non-CpG site converted Thymines are shown in green capital letters, mono-nucleotide tracks of more than 4 nucleotides appear as strike-through text and complex tetranucleotides as underlined bold text

**Cytosine-guanine dinucleotide sites (CpG):** The CpG are potentially methylated or non-methylated DNA tracks at the 5th carbon of the cytosine. To control PCR bias, one or two CpGs should be included in the primer design, as far as possible from the 3'oligonucleotide extreme. Non-methylated alleles of each CpG included in the primer will generate a mismatch near the 5' oligonucleotide extreme, thus reducing the efficiency of amplification of non-methylated allele and equilibrating the bias[11]. Including three CpG sites is allowable but it should be done cautiously because it could redirect the bias toward methylated alleles[11]. In BiCVisualizer output the CpG sites remain unconverted appearing in red and bold for easy visualization (Fig. 2).

**Thymines generated by conversion process:** After bisulfite treatment, all Cytosines from non-CpG sites will be transformed into Uracils and the subsequent PCR amplification replaces them with Thymines (Ts). In order to increase the specificity of the amplification for completely converted sequences it is important to include at least one

Table 2: Primers designed with BiCVisualizer and corresponding *in silico* specificity evaluation

| Gene | HG38 Position | F/R | Forward primer sequence / Reverse primer sequence (Annealing sequence) | Individual matches | Number of amplicons found (pairs specificity) |
|---|---|---|---|---|---|
| MBP | Chr18:77087458-77087617 | F | gagTtCGTTagTTTtagaT | 196 | 1 |
| | | R | CGAcctcaccctacaA (TtgtagggtgaggTCG) | 84 | |
| APP | Chr21: 26171148-26171257 | F | gtCGtataaaggaTtgTtgTta | 701 | 1 |
| | | R | cttcactCGttctcattctct (agagaatgagaaCGagtgaag) | 96 | |
| FMR1 | ChrX: 147911827-147911936 | F | gTCGaggggTtgagTT | 196 | 1 |
| | | R | ACTAaACGcctAactAAAA (TTtTagtTaggCGTtTagT) | 425 | |
| CAPN10 | Chr2: 240585767-240585918 | F | TCGCGgTaTtgaatggTtatagT | 631 | 1 |
| | | R | ACGctctacatcttaact~~AA~~ (TTagttaagatgtagagCGt) | 934 | |
| CLU | Chr8: 27614632-27614722 | F | agggaagaCGgggaTatTtTaT | 163 | 1 |
| | | R | GACGcctcccaAtAcc (ggTaTtgggaggCGTC) | 110 | |

Matches: Primer sequences matches of each primer in the bisulfite converted genome (mean by strand) using Bisearch/Simple-search tool as a measure of specificity, HG38 position: Genomic position of amplicon in Hg38 human genome annotation CpG sites in bold capital letters, the non-CpG site converted Thymines are shown as capital letters and complex tetranucleotides as underlined bold text

Thymine derived from a converted Cytosine, as close as possible to the 3′ end[4,14]. BiCVisualizer output colors include differential highlighting of Cs located in non-CpG sites after replacing them with Ts. A green color and bold capital letters distinguish those new Ts from the unconverted nucleotides in the original sequence (Fig. 2).

**Visualizing complexity:** Complexity sequences were emphasized in two ways: low complexity mono-nucleotide repeats and complex nucleotide sequences.

**Mono-nucleotide repeats (≥5 nt):** Mono-nucleotide repeats, defined here as mono-nucleotide tracks of five or more single nucleotide repeats become even more frequent after bisulfite conversion(for Ts or As nucleotides). Long stretches of nucleotides in primer sequence affect PCR specificity; therefore, in BiCVisualizer's output mono-nucleotide sequences are shown in cross-out letters to warn users about their presence.

**Complex tetranucleotides:** For primer designing, linguistically complex sequences would be preferred because they are less likely to be repeated in the genome, producing a specific amplification. In other words, linguistically complex nucleotides are relatively infrequent targets of unspecific amplification. In BiCVisualizer output these sequences were shown as underlined letters to facilitate their inclusion in the 3′ end of the designed oligonucleotides (Fig. 2).

**Specificity assessment:** After BiCVisualizer implementation, PCR specificity of 10 primers designed with the visual assistance of this tool was evaluated using the ePCR/BiSearch software tool to assess the number of primer pairs present and the number (or the absence) of unspecific amplicons[15]. Previous observations recommended less than 3000 primers match and the absence of *in silico* amplicons in BiSearch[5]. The absence of unspecific amplicons is indicative of total specificity[5,15].

Although BiSearch/ePCR is the only one available tool to perform the specificity assessment of this class of primers through *in silico* PCRs in the human genome, BiSearch does not assist MS-HRM primer desinging[15].

The pairs of testing primers used in this work were designed in a semi-manual manner using BiCVisualizer according to the recommendations, directed to regions of epigenetic importance from MBP, APP, FMR1, CAPN10 and CLU genes (Table 2).

## RESULTS AND DISCUSSION

In the individual specificity analysis of testing primer designed using BiCVisualizer based on *in silico* search of matches in BiSearch/ePCR tool, less than 1000 matches per primer were found in a range of 84-934 matches (mean±SD; 353.6±302) (Table 2). This range of results had clearly shown a satisfactory individual specificity measurements according to previous recommendation of less than 3000 matches in the genome per primer[5].

Regarding the primer pairs specificity assessment, only one specific amplicon and no unspecific amplicons were found for each of the five in *in silico* PCRs analyses (Table 2). This results indicated the complete specificity of the evaluated primer pairs according to the *in-silico* gold standard specificity measurement measurement[15].

The present software was developed to facilitate MS-HRM primers design following current recommendations and additionally to assist in the location of linguistically complex sequences in this process. Current recommendations for primers design became a guideline useful for PRCs in many biomedical research fields including oncology[16], neurology[17]

and pregnancy[18]. BiCVisualizer tool is available as a free easy-to-access web-tool for any scientific interested in studying DNA methylation.

Finding the high complexity sequences may be critical for PCR success, as was found in the *in silico* PCR evaluation. Sequences with known biological function such as genes, conserved regions, regulatory domains and exons have been associated with linguistic complexity[19,20] and previous software that measured this parameter were developed to address questions of biological importance[20-22]. For example, Fast PCR a commercially available software has used sequence complexity criteria for primer selection according to the alphabet capacity method[22]. However, Fast PCR is not comparable with BiCVisualizer, because it is not intended to assist MS-HRM primers design with current recommendations[22].

The present software, BiCVisualizer, is a bisulfite converted DNA visualizer for researchers interested in MS-HRM and other PCR based DNA methylation techniques. Other primer designing programs for bisulfite-based PCR techniques including Methprimer and BiSearch have different functions[15,23,24]. However, BiCVisualizer is the first available tool that allows for the visualization of CpG sites, converted Thymines, mononucleotide tracks and complex sequences at the same time, in an efficient and rapid manner. Furthermore, the primer pairs designed using BiCVisualizer primer only showed one amplicon per primer pair, indicating high specificity in the *in silico* evaluation tool[15].

Since Bisulfite-converted methylated sequences including MS-HRM, amplify less efficiently than non-methylated sequences due the lower melting temperature of non-methylated templates[8]; only if new primers designing recommendations are included[12], the increasing annealing temperature can improve the amplification of methylated templates and the decreasing annealing temperature may increases the amplification of non-methylated targets[6]. Therefore, PCR bias can be controlled in MS-HRM assays[11]. In this setting, BiCVisualizer has provided visual insights for designing HRM primers not available in other software[15,23,24].

Finally, it is recommended (as for any other kind of primer-design) to evaluate conventional primer parameters such as dimer stability and hairpin stability before experimental usage. This could be a limitation of the present work because the evaluation of the designed primers should be completed using additional software. On the other hand, this work has presented a reliable, straightforward and efficient software for primer designing providing a DNA methylation study tool for the scientific community interested in the field. The present work provides to the scientific community the BiCVisualizer, a previously unavailable software to visually assist with DNA methylation primer design including MS-HRM, one of the most used and cost-effective quantitative DNA methylation assays[5].

## CONCLUSION

BiCVisualizer software is a tool for specific locus DNA methylation primer designing, easy to use in MS-HRM. Using BiCVisualizer is possible to obtain reliable primers with satisfactory specificity in a short time.

## ACKNOWLEDGMENTS

The work of the authors was supported in part by Universidad Santo Tomás, Bucaramanga, Colombia, Universidad Nacional de Colombia, Bogotá, Colombia and the Universidad Industrial de Santander, Bucaramanga, Colombia.

## REFERENCES

1. Wong, E.M., M.C. Southey, S.B. Fox, M.A. Brown and J.G. Dowty *et al*., 2011. Constitutional methylation of the BRCA1 promoter is specifically associated with BRCA1 mutation-associated pathology in early-onset breast cancer. Cancer Prev. Res., 4: 23-33.
2. Ek, W.E., M. Rask-Andersen and A. Johansson, 2016. The role of DNA methylation in the pathogenesis of disease: What can epigenome-wide association studies tell? Epigenomics, 8: 5-7.
3. Sandoval-Hernandez, A.G., H.G. Hernandez, A. Restrepo, J.I. Munoz and G.F. Bayon *et al*., 2016. Liver X receptor agonist modifies the DNA methylation profile of synapse and neurogenesis-related genes in the triple transgenic mouse model of alzheimer's disease. J. Mol. Neurosci., 58: 243-253.
4. Wani, K. and K.D. Aldape, 2016. PCR techniques in characterizing DNA methylation. Methods Mol. Biol., 1392: 177-186.
5. Hernandez, H.G., M.Y. Tse, S.C. Pang, H. Arboleda and D.A. Forero, 2013. Optimizing methodologies for PCR-based DNA methylation analysis. Biotechniques, 55: 181-197.
6. Xiao, Z., B. Li, G. Wang, W. Zhu and Z. Wang *et al*., 2014. Validation of Methylation-Sensitive High-Resolution Melting (MS-HRM) for the detection of stool DNA methylation in colorectal neoplasms. Clin. Chim. Acta, 431: 154-163.
7. Wojdacz, T.K., A. Dobrovic and L.L. Hansen, 2008. Methylation-sensitive high-resolution melting. Nat. Protoc., 3: 1903-1908.

8. Hussmann, D. and L.L. Hansen, 2018. Methylation-Sensitive High Resolution Melting (MS-HRM). Methods Mol. Biol., 1708: 551-571.
9. Hernandez, H.G., M.F. Mahecha, A. Mejia, H. Arboleda and D.A. Forero, 2014. Global long interspersed nuclear element 1 DNA methylation in a Colombian sample of patients with late-onset Alzheimer's disease. Am. J. Alzheimer's Dis. Other Dementias, 29: 50-53.
10. Zhao, H.B.X. and F.G. Jiang, 2015. Value of PAX1 methylation analysis by MS-HRM in the triage of atypical squamous cells of undetermined significance. Asian Pac. J. Cancer Prev., 16: 5843-5846.
11. Wojdacz, T.K., T. Borgbo and L.L. Hansen, 2009. Primer design versus PCR bias in methylation independent PCR amplifications. Epigenetics, 4: 231-234.
12. Wojdacz, T.K., L.L. Hansen and A. Dobrovic, 2008. A new approach to primer design for the control of PCR bias in methylation studies. BMC Res. Notes, Vol. 1, No. 1. 10.1186/1756-0500-1-54.
13. Xi, Y. and W. Li, 2009. BSMAP: Whole genome bisulfite sequence MAPping program. BMC Bioinform., Vol. 10, No. 1. 10.1186/1471-2105-10-232.
14. Wojdacz, T.K., J.A. Windelov, B.B. Thestrup, T.E. Damsgaard, J. Overgaard and L.L. Hansen, 2014. Identification and characterization of locus-specific methylation patterns within novel loci undergoing hypermethylation during breast cancer pathogenesis. Breast Cancer Res., Vol. 16, No. 1. 10.1186/bcr3612.
15. Aranyi, T. and G.E. Tusnady, 2007. BiSearch: ePCR Tool for Native or Bisulfite-Treated Genomic Template. In: PCR Primer Design, Yuryev, A. (Ed.). Chapter 20, Springer, New York, USA., ISBN-13: 9781588297259, pp: 385-402.
16. Daugaard, I., D. Hussmann, L. Kristensen, T. Kristensen and T.E. Kjeldsen *et al.*, 2018. Chronic lymphocytic leukemia patients with heterogeneously or fully methylated LPL promotor display longer time to treatment. Epigenomics 10: 1155-1166.
17. Coppede, F., P. Tannorella, A. Stoccoro, L. Chico, G. Siciliano, U. Bonuccelli and L. Migliore, 2017. Methylation analysis of DNA repair genes in Alzheimer's disease. Mechan. Ageing Dev., 161: 105-111.
18. Chavez-Lizarraga, D., P. Zarate-Segura, J. Beltran-Montoya, C. Canchola-Sotelo, F. Vadillo-Ortega and E. Chavira-Suarez, 2018. DNA methylation variability in a single locus of the RXRα promoter from umbilical vein blood at term pregnancy. Biochem. Genet., 56: 210-224.
19. Monge, R.E. and J.L. Crespo, 2015. Analysis of data complexity in human DNA for gene-containing zone prediction. Entropy, 17: 1673-1689.
20. Pirhaji, L., M. Kargar, A. Sheari, H. Poormohammadi, M. Sadeghi, H. Pezeshk and C. Eslahchi, 2008. The performances of the chi-square tesmeasures for signal recognition in biological sequences. J. Theor. Biol., 251: 380-387.
21. Kalendar, R., D. Lee and A.H. Schulman, 2011. Java web tools for PCR, *in silico* PCR and oligonucleotide assembly and analysis. Genomics, 98: 137-144.
22. Kalendar, R., D. Lee and A.H. Schulman, 2014. Fast PCR Software for PCR, *In silico* PCR and Oligonucleotide Assembly and Analysis. In: DNA Cloning and Assembly Methods, Valla, S. and R. Lale (Eds.). Chapter 18, Springer, New York, USA., ISBN-13: 9781627037631, pp: 271-302.
23. Tusnady, G.E., I. Simon, A. Varadi and T. Aranyi, 2005. BiSearch: Primer-design and search tool for PCR on bisulfite-treated genomes. Nucl. Acids Res., Vol. 33, No. 1. 10.1093/nar/gni012.
24. Li, L.C. and R. Dahiya, 2002. MethPrimer: Designing primers for methylation PCRs. Bioinformatics, 18: 1427-1431.