FEBS Letters 584 (2010) 739–744

FEBS Letters

journal homepage: www.FEBSLetters.org

Methyl-Typing: An improved and visualized COBRA software for epigenomic studies

Cheng-Hong Yang [a,b], Li-Yeh Chuang [c,*], Yu-Huei Cheng [a], De-Leung Gu [f], Chung-Ho Chen [d,e], Hsueh-Wei Chang [f,g,h,*]

[a] Department of Electronic Engineering, National Kaohsiung University of Applied Sciences, Kaohsiung, Taiwan
[b] Department of Network Systems, Toko University, Chiayi, Taiwan
[c] Department of Chemical Engineering, I-Shou University, Kaohsiung, Taiwan
[d] Department of Dentistry, College of Dental Medicine, Kaohsiung Medical University, Kaohsiung, Taiwan
[e] Department of Oral and Maxillofacial Surgery, Kaohsiung Medical University Hospital, Kaohsiung, Taiwan
[f] Department of Biomedical Science and Environmental Biology, Kaohsiung Medical University, Kaohsiung, Taiwan
[g] Graduate Institute of Natural Products, Kaohsiung Medical University, Kaohsiung, Taiwan
[h] Center of Excellence for Environmental Medicine, Kaohsiung Medical University, Kaohsiung, Taiwan

ARTICLE INFO

Article history:
Received 14 October 2009
Revised 14 December 2009
Accepted 16 December 2009
Available online 22 December 2009

Edited by Paul Bertone

Keywords:
Methylation
COBRA
Restriction enzyme
Software
Promoter

ABSTRACT

Combined bisulfite restriction analysis (COBRA) is one of the most commonly used methylation quantification methods. However, it focuses on relatively few restriction enzymes. Here, we present Methyl-Typing, a web-based software that provides restriction enzyme mining data for methyl-cytosine-containing sequences following bisulfite-conversion. Gene names, accession numbers, sequences, PCR primers, and file upload are accessible for input. Promoter sequences and restriction enzymes for CpG- and GpC-containing recognition sites are retrieved. Four representative enzymes were tested successfully by COBRA on the experimental work. Therefore, the Methyl-Typing tool provides a comprehensive COBRA-restriction enzyme mining. It is freely available at http://bio.kuas.edu.tw/methyl-typing.

© 2009 Federation of European Biochemical Societies. Published by Elsevier B.V. All rights reserved.

1. Introduction

Currently, combined bisulfite restriction analysis (COBRA) [1] is one of the most commonly used methylation methods in laboratories [2]. In principle, a technique using any kind of restriction enzymes to distinguish between the methylated- and the unmethylated-sequences with bisulfite-conversion are regarded as the COBRA method. However, the traditional COBRA approach only relies on a few restriction enzymes, such as BstUI (5′-CG$^{\downarrow}$CG-3′) [2] and Taq$^{\alpha}$I (5′-T$^{\downarrow}$CGA-3′) [1]. Other restriction enzymes available for COBRA are less frequently mentioned, such as HinP1I (5′-G$^{\downarrow}$CGC-3′), HpyCH4IV (5′-A$^{\downarrow}$CGT-3′), and AciI (5′-G$^{\downarrow}$CGG-3′). This may in part be due to the fact that restriction enzyme mining tools for possible methylation sequences are poorly developed. Moreover, the traditional COBRA approach is not naturally specific to CpG islands in the promoter region but it depends on the user-defined sequence. Therefore, the integration for COBRA-restriction enzyme mining, CpG island searching, and promoter prediction is still challenging.

In order to circumvent this constraint, we have developed a novel visualization software, named Methyl-Typing, which provides comprehensive restriction enzymes for methyl-cytosine-containing sequences after bisulfite-conversion, i.e. unmethylated cytosine converts to uracil (regarded as thymine for PCR amplification) while 5-methylcytosine remains unchanged. Moreover, the insulators' CCCTC-binding factor (CTCF)-binding site database (CTCFBSDB) [3] is implemented in Methyl-Typing. The insulators of chromatin, such as CTCF, can block the activity of a down-stream enhancer and are neutralized by methylation [4], thereby contributing to gene regulation. In conclusion, Methyl-Typing is a fast and efficient tool for providing all possible methylation sites of restriction enzymes.

Abbreviations: COBRA, combined bisulfite restriction analysis; CTCF, CCCTC-binding factor; CTCFBSDB, CCCTC-binding factor-binding site database; RFLP, random fragment length polymorphism

* Corresponding authors. Address: Department of Biomedical Science and Environmental Biology, Kaohsiung Medical University, Kaohsiung, Taiwan (H.-W. Chang). Fax: +886 7 312 5339.
E-mail addresses: chuang@isu.edu.tw (L.-Y. Chuang), changhw@kmu.edu.tw (H.-W. Chang).

0014-5793/$36.00 © 2009 Federation of European Biochemical Societies. Published by Elsevier B.V. All rights reserved.
doi:10.1016/j.febslet.2009.12.026

2. Materials and methods

2.1. Computational analysis – implementation

Methyl-Typing, a web-based software package, was designed and implemented under JSP (Java Server Page) and Servlet language. Sequences were retrieved online from GenBank and in silico PCR of UCSC Genome Browser [5] when the accession numbers and primers were used for input, respectively. A promoter database for human and mouse genomes, and a restriction enzyme database were downloaded from DBTSS [6] and REBASE version 806 [7], respectively. CpG island mining was derived from the web site for CpG island search [8]. The CTCF-binding site database for vertebrate genomic insulators was hyperlinked and fed to CTCFBSDB [3] for online analysis. These databases and online retrieved websites will be checked and updated every 2 months. The bisulfite sequencing primers (BSP) were designed in Methyl-Typing according to the criteria of user manual for the freeware-Methyl Primer Express® Software v1.0 (www.appliedbiosystems.com/methylprimerexpress). Subsequently, COBRA analysis is able to be performed using the BSP primers for PCR amplification and the mined COBRA enzymes for digestion.

For bisulfite T stretch, all the C nucleotides in a sequence are converted to T nucleotides by default (computer-generated bisulfite-converted sequences) and the continuous T nucleotides are processed in order to be visualized.

Three kinds of functions are incorporated in the Methyl-Typing system: the modules for input, analysis, and output. A demonstration and the user manual for Methyl-Typing are available for download at http://bio.kuas.edu.tw/methyl-typing and http://bio.kuas.edu.tw/methyl-typing/userManual_info.jsp, respectively. Some videos are provided on the homepage for tutorial purpose.

2.2. Experimental work – sample collection and DNA extraction

Four oral tumors (Nos. 1–4) and four normal gingiva (Nos. 5–8) tissue samples were collected with the approval of the institutional review board at Kaohsiung Medical University, Taiwan. Tissue DNA was extracted with the DNeasy tissue kit (Qiagen, Valencia, CA, USA), as previously described [9].

2.3. Experimental work – bisulfite modification

Bisulfite-conversion was performed with 1 μg of genomic DNA for the EpiTect® Bisulfite kit (Qiagen) according to the manufacturer's instructions. The purified bisulfite-modified DNA was twice eluted in 20 μl buffer EB.

2.4. Experimental work – amplification of bisulfite-converted DNA and restriction enzyme digestion (COBRA)

Bisulfite-converted DNA samples were added to the PCR reaction mixture (10 μl) containing 1 μl of 10X PCR buffer, 0.3 μl of 50 mM $MgCl_2$, 0.2 μl of 10 mM dNTPs, 0.6 μl of DMSO, 0.14 μl of 5U Taq enzyme, 0.12 μl of 350 μg/ml primers mix, and 7.64 μl of DNA in water. Two primer sets specific for the promoter of the TEKT4 gene (P1**C**-F1: 5′-TAGGAAG**C**TtTTGGTTTAtAtAGTG-3′, P1**T**-F1: 5′-ATAGGAAG**t**TtTTGGTTTAtAtAGTGT-3′, P1**C**-R1: 5′-AACCACTCCTCCA**G**CAaaTAC-3′, and P1**T**-R1: 5′-aAACCACTCCTCCA**a**CAaaTAC-3′) were designed to improve the annealing for PCR concerning the presence and absence of 5′-methylation of the **C** nucleotides in the target sequence. Lowercase letters in primer sequences represent the bisulfite-conversion of unmethylated C (target sequences) to t (primer sequences). Bold and underlined letters listed in primer sequences represent the possible **t/C** polymorphism (reverse in **a/G**) of primer sequences which is complementary to the target sequence with and without conversion after bisulfite treatment, respectively. The touch-down program was performed as described previously [10]. PCR amplicons were digested with restriction enzymes (Bstu1, HinP1I, HpyCH4IV, and HaeIII) mining by Methyl-Typing and run in 4% agarose gel electrophoresis. The methylation results of the COBRA assay was validated by bisulfite sequencing [11].

3. Results

3.1. Computational analysis – input module

The following modes of input are available in the module: the multiple gene name input in official gene symbol, multiple accession number input for nucleotide sequence retrieval, multiple template sequence input using free format, primer input, and file upload (Fig. 1A–E). Accession number inputs are available because Methyl-Typing is capable of online retrieval of the corresponding sequences. In sequence input, FASTA template sequences (i.e. >name [first line]-sequence [second-line], with A, T, C, and G) are accepted. Upper and lower case is not significant and all other characters, including spaces and digits, are ignored. Regular PCR primer sequences in any orientations, rather than the methylation-specific primers, are acceptable to process to in silico PCR to generate sequence for further analysis. A sequence file is also acceptable for uploading in text format.

3.2. Computational analysis – analysis module

For gene name input (Fig. 1A), the human and mouse promoter sequences from DBTSS [6] can be searched. After gene name input, the promoter sequence is retrieved (as shown in the user manual). For input formats other than the gene name input, promoter retrieval can not be performed due to the nature of the sequence. Subsequently, the CpG island search is processed with adjustable %GC, observed CpG/Expected CpG, and length of CpG islands. The result of the CpG island search and restriction enzyme mining for random fragment length polymorphism (RFLP) analysis to putative methylated sequences are provided. Visualizations for the CpG mapping, CpG distribution, and the CpG sequence are provided as well (as shown in the user manual). Methyl-Typing also provides a CTCFBS search [3] and bisulfite T stretch analysis (described later in Fig. 2F and 2G).

3.3. Computational analysis – output module

In the example of TEKT4 gene input, the result of the methylation-RFLP analysis is shown in Fig. 2. In Fig. 2A, four functions are provided: sequence viewing, visualization, CTCFBS search, and bisulfite T stretch. In Fig. 2B, the TEKT4 promoter sequence is provided. The original or bisulfite-treated/methylated sequences with CpG islands and the mined restriction enzymes are colored in green and red, respectively. The bisulfite-treated/unmethylated sequence is also provided in parallel for comparison (not shown). By clicking the BSP box as shown in Fig. 2C, the suitable primers required for PCR in COBRA assay are designed as well as their exact restriction fragment lengths (not shown). In Fig. 2C, all the mined enzymes are provided in a chart by clicking the boxes of sequence and visualization as shown in Fig. 2A. In addition to visualization, Methyl-Typing also provides an interactive interface for selected enzymes in CpG islands (Fig. 2C–E). In the example of sequence 5′-CGCG-3′ (the recognition site for AccII, Bsh1236, BspFNI, BstFNI, BspUI, and MvnI), the CpG island distribution is provided by moving the mouse-over the bars for restriction enzyme sites (Fig. 2D).

Download English Version:

https://daneshyari.com/en/article/2048197

Download Persian Version:

https://daneshyari.com/article/2048197

Daneshyari.com