Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

A simple and cost-effective approach for technical validation of next generation methylation sequencing data

Ali Javadmanesh
Ferdowsi University of Mashhad
Afsaneh Mojtabanezhad Shariatpanahi
Reza Radiotherapy and Oncology Center
Ehsan Shams Davodly
Reza Radiotherapy and Oncology Center
Marjan Azghandi
Reza Radiotherapy and Oncology Center
Maryam Yassi
Reza Radiotherapy and Oncology Center
Mehdi Heidari
Reza Radiotherapy and Oncology Center
Matin Kerachian
McGill University Faculty of Medicine
Mohammad Amin Kerachian (amin.kerachian@mail.mcgill.ca)
Mashhad University of Medical Sciences https://orcid.org/0000-0001-5016-2882

Research article

Keywords: DNA Methylation Sequencing, Validation, High resolution melting, Bisulfite

Posted Date: September 9th, 2019

DOI: https://doi.org/10.21203/rs.2.14216/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Abstract

Background DNA methylation is a fundamental epigenetic process that, in most cases, modulates genetic expression levels. Changes in DNA methylation, either hypo- or hypermethylation, have a key role in many biological processes and several human diseases such as cancer. In the current study, we offered an approach to validate the next generation methylation sequencing data.

Methods Genomic DNA was extracted from target and control samples (6 in each group), followed by bisulfite conversion. Next generation methylation sequencing and methylation sensitive high-resolution melting assay were carried out. The primers for methylation sequencing validation were designed by R programming language.

Results In the current study, two groups, case and control, were discriminated based on methylation sequencing results and the real time PCR-based results were in accordance with the next generation methylation sequencing.

Discussion Methylation sensitive high-resolution melting validation assay is a simple and cost-effective method, which confirmed next generation methylation sequencing results.

Background

DNA methylation is one of many fundamental epigenetic processes that modulates the expression levels of genes [1]. It functions as an annotation system for marking the genetic text, thus providing instruction as to how and when to read the information and control transcription. Molecularly speaking, DNA methylation is a covalent modification that occurs exclusively on cytosine nucleotides [2]. In vertebrates, it is characterized by the addition of a methyl or hydroxymethyl group on the C5 position of cytosine. These modifications are most commonly associated with CpG site, that being sites where a cytosine is immediately followed by a guanine in a 5’ to 3’ direction. Non-CpG methylation in a CHH and CHG context (where H = A, C or T) can be found in embryonic stem cells and in plants [3]. Changes in DNA methylation, either hypo- or hypermethylation, play a key role in many biological processes and several human pathologies [4]. DNA methylation is also associated to the process of gene imprinting which has been linked to syndromes such as Prader-Willi, Angelman, or Beckwith Wiedemann. In cancer cells, it is likely that a great deal of changes in tumor cells have potential epigenetic origins [2]. Abnormal DNA methylation in cancer cells may be classified into two categories: site specific CpG island promoter hypermethylation and global DNA hypomethylation [1]. With the use of high-throughput technologies when detecting cancer mutations, it has been demonstrated that many of the affected genes are involved in DNA methylation metabolism or in the control of chromatin structure. These results increase the possibility that the epigenetic state of tumors may reflect one of the many mutational consequences that occur in cancer development [5]. It should be noted that aberrant methylation appears to take place prior to transformation of cancer cells. Thus, aberrant methylation could potentially be used as an early detection approach [6]. In the past few years, the discovery of non-invasive cancer detection assays based on aberrant methylation states has increased in demand and

interest [7]. In addition to high-throughput mutation detection techniques, other methods for methylation detection have emerged for methylation cancer study and methylation marker discovery. These techniques can be divided into three categories: 1) methylation content assay: high-performance capillary electrophoresis or high-performance liquid chromatography; 2) methylation pattern and profiling: restriction landmark genomic scanning, methylated CpG-island amplification, amplification of inter-methylated sites; 3) candidate gene approach: methylation-sensitive restriction endonuclease-PCR/Southern, bisulfite sequencing, methylation-specific PCR, MethyLight technology [8].
In any methylation marker discovery or cancer study, it is critical to validate the methyl sequencing results conducted by next generation sequencing (NGS). Most commonly, generating NGS results requires the design of complementary nucleic acids as site-specific markers that target the genome. In many cases, the process of designing the primers and probes can be laborious and may not properly represent the aberrant methylation alterations.
In the current study, we offered a novel, cost-effective and simple approach to validate the NGS-based methylation results proposing the use of a methylation sensitive high-resolution melting assay.

Methods

Sample collection and preparation

Cancerous tissue samples from patients with colorectal cancer (N=6) confirmed by two gastroenterology pathology experts were considered as target or case tissues (assigned as cases: T20, T31, T35, T45, T65, T67). Normal controls tissues (N=6) were from individuals who underwent colonoscopy screening that were negative for either adenomatous polyps or cancer through the entire colon (assigned as controls: N4, N7, N8, N10, N14, N16). Demographic characteristics, colonoscopy reports, history of drug intake, smoking, as well as medical history were all collected. The case and control groups were matched by their demographic features. The protocol was approved by Mashhad University of Medical Sciences (Grant number: 961906). Informed written consent had been obtained from all participants in this study and have the participants permission to get published.

Nucleic Acid Isolation from Tissue

Genomic DNA was extracted from 5-25 mg of fresh tissues using QIAamp® Fast DNA Tissue Kit (Qiagen, Valencia, CA, USA) following the manufacturers instruction. The extracted DNA was evaluated by Epoch Microplate Spectrophotometer (Winooski, Vermont, USA).

Bisulfite Treatment

Two micrograms of extracted DNA undergone sodium bisulfite conversion and DNA recovery using the EpiTect Fast Bisulfite Conversion Kits (Qiagen, Germany) according to the manufacturer’s instructions. For bisulfite conversion of genomic DNA, a number of

different commercial kits for bisulfite conversion are available. The kit and protocol that best suits the planned experiment was selected accordingly.

Methylation Next Generation Sequencing and Analysis

The global methylation profiles of target and normal control tissues were analyzed using the SureSelect Human Methyl-seq. This platform assesses 84 mega base (MB) of genome, 3.7 million CpGs, 19.6 Mb CpG Islands, 37 MB Gencode promoters, 48 MB enhancers, CpG island shores/shelves ±4 and DNase I hypersensitive sites. The protocol is based on biotinylated RNA probes to capture target epigenomic regions and only requires 2–3 µg of gDNA for library construction. The gDNA is randomly fragmented to mean sizes of approximately 250 bp, then fragments are ‘‘fixed,’’ creating blunt ends and adding dA in the 3’ ends so the methylated modified adapters bind to these fragments [9]. Following bisulfite treatment, sequencing was performed on the Illumina HiSeq 4000 sequencer (San Diego, USA).

The Human Methyl-seq analysis was composed of three steps in the preprocessing stage before detecting differentially methylated regions. Firstly, the total reads were assessed by Quality Control tool [10]. Secondly, the raw sequencing reads were cleaned by trim galore [11]. Thirdly, the raw bisulfite sequencing data were converted into a number of methylated reads and covered reads of cytosines (including unmethylated/methylated reads) by aligning them to the human reference genome (GRCh37/19) using the Bismark [12].

Target Identification and Primer Design

We presented a primer design for methylation sequencing validation by R programming language. Methylation independent primers were designed based on the following criteria:

1. The amplification length is between 98-170 base pairs.
2. Only one CpG site is located in this length.
3. Coverage is more than the average depth per CpG sites in data.

The framework of our method is shown in Algorithm 1.

Methylation Quantitation assay

Methylation analysis was performed by methylation sensitive high-resolution melting (MS-HRM). The MS-HRM protocol consists of PCR amplification of bisulfite-modified DNA. Two primers sets, A and B, were used to amplify bisulfite-treated DNA A-F: 5′-TGAGAGGGTTTTGTGGGTTATT-3′, A-R: 5′-TCCCCTAAACTCAAACTATCACTC-3′, and B-F: 5′-TAAGATTTAGAGGAAAATGTTTGTT-3′, B-R: 5′-ACCAAACCAAAACACTATACTACAT-3′ were designed to amplify both methylated and unmethylated bisulfite-treated DNA that did not amplify unmodified genomic DNA. Polymerase chain reaction (PCR) amplification and HRM analysis were carried out sequentially on a light Cycler® 96 System (Roche, Switzerland). PCR reactions were carried out in triplicates in a 10 µl total volume using HOT FIREPol EvaGreen Green Master Mix (Solis BioDyne, Estonia), consisting of 300 nM of each primer, 0.5 µl of bisulfite modified template. The amplification run was 15 min at 95 °C, followed by 50 cycles of 20 s 95 °C, 30s at the primer annealing temperature (55 °C) and 30s at 72 °C. HRM analyses were performed at the temperature ramping from 65 to 97 °C. Fluorescence acquisition setting was carried out at temperature recommended by the manufacturer. The melting curves and melting peak were normalized by calculation of the ‘line of best fit’ in between two normalization regions before and after the major fluorescence decrease representing the melting of the PCR product using the software version 1.1 provided with the LightCycler® 96 System.

Statistical analysis

To compare characteristics of the different groups of target and control samples, two sample *t* test was performed. The p value less than 0.05 was considered statistically significant.

Results And Discussion

In the current study, two groups, case (T20, T31, T35, T45, T65, T67) and control (N4, N7, N8, N10, N14, N16) were discriminated based on methylation sequencing results. Their NGS features are indicated in Table 1. In order to validate the next generation methylation sequencing results, two A and B primer sets, were used to target two different regions on the bisulfite-modified DNA. The primers designed by R programming where shown in figure 1.

Methylation sensitive high-resolution melting assay was conducted with LightCycler® 96 System and their results were represented in figure 2. The real time PCR results were in accordance with the next generation methylation sequencing. There was a significantly shift (p value <0.05) for the set primer A in methylated and unmethylated, which were control and case samples, respectively. Besides, for the set primer B, methylated (control) versus unmethylated (case) samples displayed an extra peak in their melting. Only one case sample (T65) had three triplicate distinctive peaks in its melting peaks, which could be explained by the small percentage of methylation (13.72%) in this sample that is zero or nearly zero in other case samples.

Several other methods have been reported to technically validate next generation methylation sequencing data such as quantitative methylation specific PCR [13, 14] which their primers and probes could be sometimes laborious to design and mostly expensive.

Our study has several strengths. First, it is a simple and cost-effective assay. Second, the R programming algorithm could be used to identify significant primer regions for methylation sensitive high-resolution melting assay, in similar studies. Study limitations should also be noted; for samples with a small fraction of methylation the results might be variable and inconsistent.

Conclusions

We have presented a strategy to validate technically methyl NGS results based on MS-HRM in this study. This approach is a simple and cost-effective method, which confirms properly methyl NGS results.

Abbreviations

MB: Mega base

NGS: Next generation sequencing

MS-HRM: Methylation sensitive high-resolution melting

PCR: Polymerase chain reaction

Declarations

Acknowledgements

Our sincere thanks also go to Mr. Ebrahim Pouladin and Mrs. Nafiseh Shalchi for their close support in colorectal cancer research programs. We would also give special thanks to

Dr. Abdorasoul Hayatbakhsh and Dr Ghodratollah Soltani for their cooperation in this study.

Funding

This study was supported financially by Reza Radiotherapy and Oncology Center, Mashhad University of Medical Sciences (Grant number: 961906) and Iran National Science Foundation (INSF; proposal number: 93048371).

Availability of data and materials

Not applicable.

Authors' contributions

AJ, AMS, ESD, MA, MY, MH, MK, MAK were all participated in study design, data analysis, and preparation of the drafted manuscript. All authors read and approved the final manuscript.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Conflict of interest

The authors declare no conflict of interest, financial or otherwise.

References

1. 1. Kerachian MA, Kerachian M. Long interspersed nucleotide element-1 (LINE-1) methylation in colorectal cancer. Clinica Chimica Acta. 2018.
2. 2. Dor Y, Cedar H. Principles of DNA methylation and their implications for biology and medicine. The Lancet. 2018.
3. 3. Van Wesenbeeck L, Janssens L, Meeuws H, Lagatie O, Stuyver L. Droplet digital PCR is an accurate method to assess methylation status on FFPE samples. Epigenetics. 2018;13(3):207-13.
4. 4. Yassi M, Davodly ES, Shariatpanahi AM, Heidari M, Dayyani M, Heravi-Moussavi A, et al. DMRFusion: A differentially methylated region detection tool based on the ranked fusion method. Genomics. 2018.
5. 5. Klutstein M, Nejman D, Greenfield R, Cedar H. DNA methylation in cancer and aging. Cancer research. 2016;76(12):3446-50.
6. 6. Shariatpanahi AM, Yassi M, Nouraie M, Sahebkar A, Tabrizi FV, Kerachian MA. The importance of stool DNA methylation in colorectal cancer diagnosis: A meta-analysis. PloS one. 2018;13(7):e0200735.
7. 7. Rokni P, Shariatpanahi AM, Sakhinia E, Kerachian MA. BMP3 promoter hypermethylation in plasma-derived cell-free DNA in colorectal cancer patients. Genes & genomics. 2018;40(4):423-8.
8. 8. Ma M, Zhu H, Zhang C, Sun X, Gao X, Chen G. "Liquid biopsy"—ctDNA detection with great potential and challenges. Annals of translational medicine. 2015;3(16).9.
9. Soto J, Rodriguez-Antolin C, Vallespin E, De Castro Carpeno J, De Caceres II. The impact of next-generation sequencing on the DNA methylation–based translational cancer research. Translational Research. 2016;169:1-18. e1.
10. Wingett SW, Andrews S. FastQ Screen: A tool for multi-genome mapping and quality control. Version 2. F1000Res. 2018; 7: 1338.
11. Babraham Bioinformatics Trim Galore. https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/. Accessed 6 Oct 2017.
12. Krueger F, Andrews SR: Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications.*Bioinformatics.* 2011;27(11):1571–1572.
13. 1 Kisiel JB, Dukek BA, Kanipakam RV, Ghoz HM, Yab TC, Berger CK, et al. Hepatocellular Carcinoma Detection by Plasma Methylated DNA: Discovery, Phase I Pilot, and Phase II Clinical Validation. Hepatology. 2018.
14. 1 Anderson BW, Suh Y-S, Choi B, Lee H-J, Yab TC, Taylor W, et al. Detection of Gastric Cancer with Novel Methylated DNA Markers: Discovery, Tissue Validation, and Pilot Testing in Plasma. Clinical Cancer Research. 2018:clincanres. 3364.2017.

Table

Table 1: Next Generation Methylation Sequencing Stats for Studied Samples												
	Primer A: Chr 20 , Location of CpG site: 58660982 product length = 157 bp											
Type of group	T*20	T31	T35	T45	T65	T67	N**4	N7	N8	N10	N14	N16
Presentage of methylation	8.57	32	31.57	50	37.5	40.54	68.42	72	84.37	96.66	81.48	78.78
Coverage	35	25	38	46	40	37	38	25	32	30	27	33
	Primer B: Chr 21 , Location of CpG site: 48090630 product length = 100 bp											
Type of group	T20	T31	T35	T45	T65	T67	N4	N7	N8	N10	N14	N16
Presentage of methylation	1.72	0	0	0	13.72	0	50	58.53	69.84	43.75	47.16	40.9
Coverage	58	45	57	52	51	57	68	82	63	48	53	44

*T is for target or case sample

** N is for control sample

Algorithm 1

Due to technical limitations, Algorithm 1 is available for download in the supplementary section.

Figures

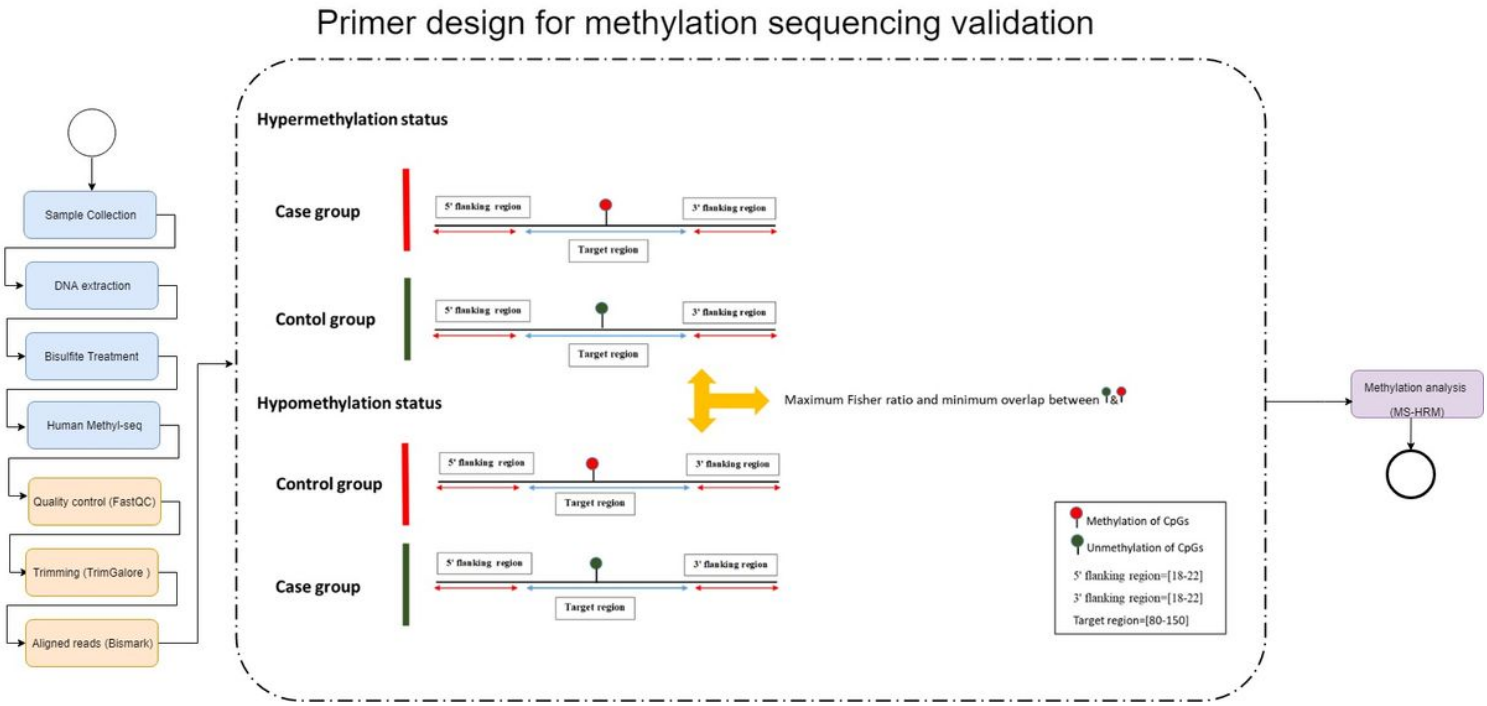


Figure 1

Structure of primer design. Solid spots are methylated (red) or unmethylated (green) of CpG sites located in approximately the middle of the target region, between 3’ and 5’ flanking regions.

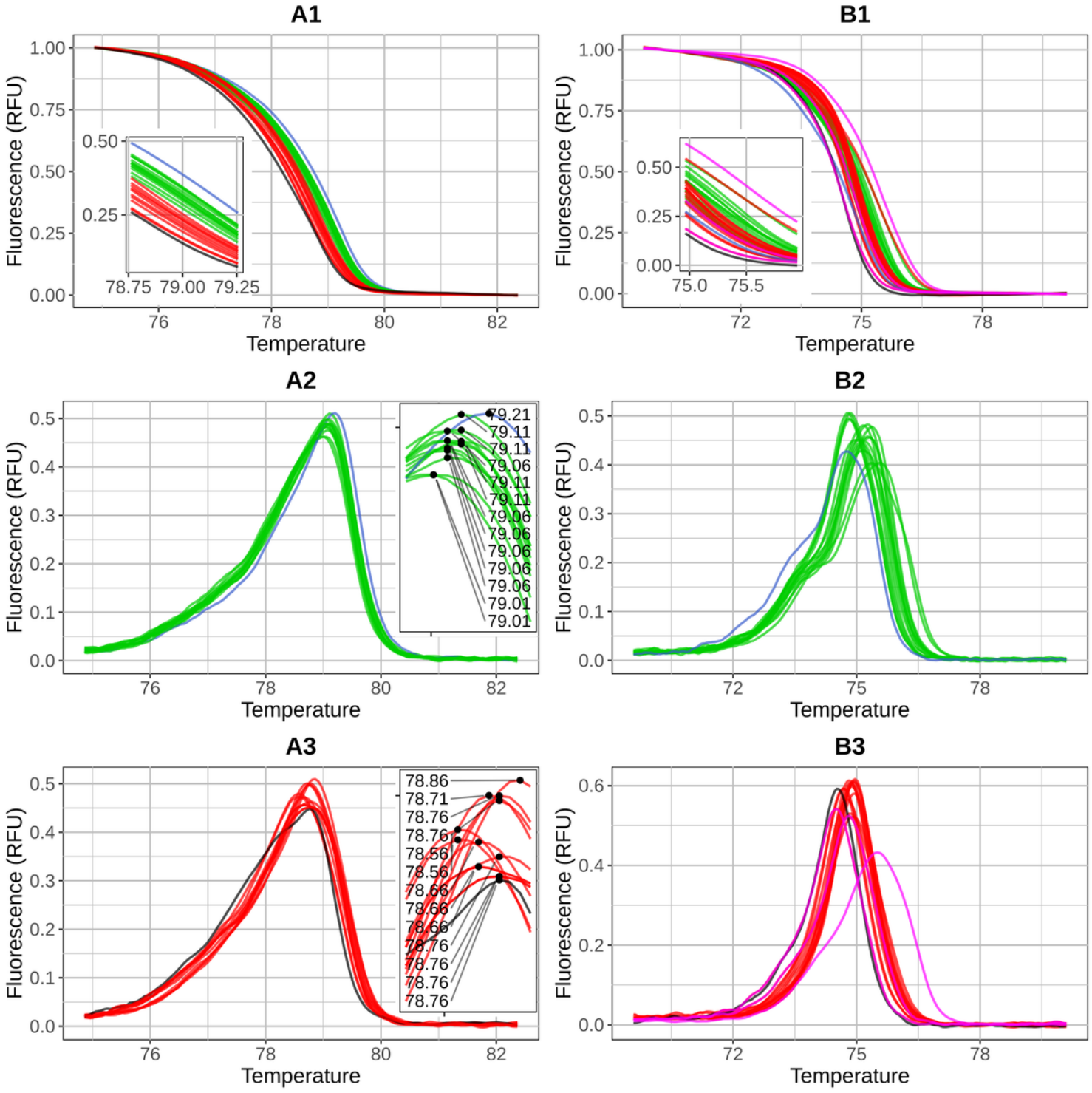


Figure 2

Normalized melting curves and normalized melting peaks of case tissues (red line) and control tissues (green line). Melting curves (A1, B1) are generated by graphing fluorescence against temperature. Melting peaks (A2, A3 and B2, B3) are generated by taking the negative derivative of fluorescence with respect to temperature. Methylated bisulfite converted control DNA and unmethylated bisulfite converted control

DNA (EpiTect Control DNA – Qiagen) shown in blue and black, respectively. The case sample (T65) is shown in pink in B1 and B3.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- algor1.jpg