# The Amount of Bisulfite-converted DNA for Reliable Age Prediction


**Sae Rom Hong[1] and Kyoung-Jin Shin[1,2]**

*[1]Department of Forensic Medicine, Yonsei University College of Medicine, Seoul, Korea*
*[2]Graduate School of Medical Science, Brain Korea 21 Project, Yonsei University, Seoul, Korea*


## Introduction

Recent studies concerning age prediction based on DNA methylation (DNAm) have substantially improved the stability and reliability of prediction results, for further usage as investigation leads. In forensics, PCR amplification of chemically converted or enzymatic deaminated DNA is a prerequisite for most existing DNAm analysis platforms and age prediction models based on those results. The age-correlated, c ontinuous trait of DNAm underlies many DNA molecules employed in PCR, since the DNAm status of a locus on one strand only can present as 0 (unmethylated) or 1 (methylated). Therefore, the amount of DNA for the PCR step strongly influences the detected DNAm values. However, previous studies only considered the amount of genomic DNA, not modified DNA. In this study, we quantified the bisulfite-con verted EpiTect PCR Control DNA and 20 human blood DNA samples using BisQuE, which is a quantitative real-time PCR system to measure the amount, conversion efficiency, and degradation level of bisulfit e-converted DNA (BS-DNA). To find the differences in DNAm levels according to the amount of BS-DNA for PCR, 10 ng and 1 ng of reference BS-DNA samples and the blood-derived BS-DNA samples were exp loited for target amplicon-based massively parallel sequencing (MPS). This MPS panel consists of nine age-associated markers on the *ELOVL2, FHL2, KLF14, MIR29B2CHG, TRIM59, EDARADD, ASPA, NHLRC1,* and *PDE4C* genes. We obtained the DNAm data through BSBolt and R, and statistical analyses for determining the differences of the DNAm were performed using R and EXCEL. These results can suggest a rel iable range for the amount of DNA for each marker when predicting age, and would offer the error range of the detected DNAm value when grappling with a low amount of DNA.

## Materials and Methods

### *Sample preparation*

- Converted with EZ DNA Methylation-Gold (Zymo) and quantified using BisQuE, which is suggested by Hong and Shin [1]
- DNA methylation control *(a total of 36 samples)*
  - Human Methylated and Non-Methylated (WGA) DNA Set (Zymo)
  - Mixed for 0%, 5%, 10%, 25%, 50%, 75%, 90%, 95%, and 100% methylated status
  - Duplication of both 10 ng and 1 ng (based on the long-sized amplicon results from [1])
- Blood-deriven human DNA samples *(a total of 100 samples)*
  - 20 Korean people aged 20-74 (10 males and 10 females)
  - 10 ng, 2 ng, 1 ng, 500 pg, and 250 pg (based on the long-sized amplicon results from [1])

### *Age-associated in house MPS panel*

- Simultaneously amplified CpG markers on the *ELOVL2, FHL2, KLF14, MIR29B2CHG, TRIM59, EDARADD, ASPA, NHLRC1*, and *PDE4C* genes

### *MPS library preparation and run*

- The two-round PCR step for MPS library preparation using unique dual index primer pairs (NEB)
- Quantified using KAPA library quantification kits (KAPA Biosystems) and 2100 Bioanalyzer (Agilent), and normalized to 10nM and then pooled in equal volumes
- Sequenced on the MiSeq System (Illumina) using MiSeq Reagent Kit V2 (500-cycles)

### *MPS data analysis*

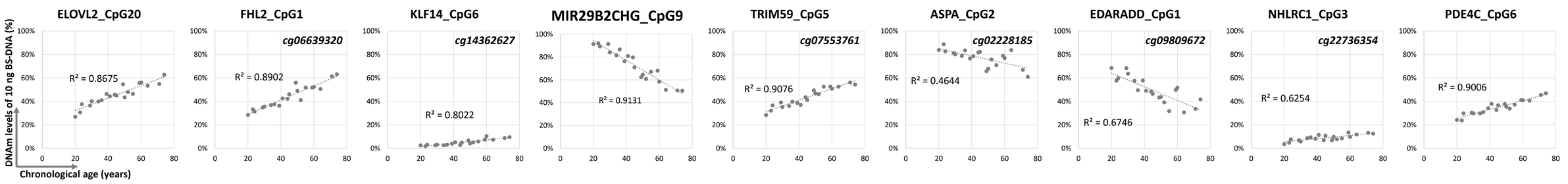
- Trim Galore V0.6.6: Trimmed adaptors in FASTQ files
- BiSulfite Bolt (BSBolt) v1.4.8
  - Alignment based on BWA-MEM using *in silico* BS-converted genomic reference sequences
  - Methylation calling with SAMtools: both base quality and mapping quality > 30
  - Methylation matrix generation and BS conversion ratio
- Microsoft EXCEL and R

## Results

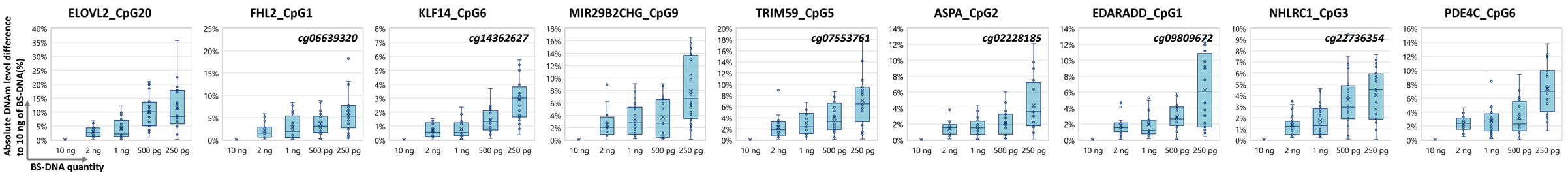
### *Measured DNAm levels in control BS-DNA*


- The average DNAm levels of each CpG site in amplicons were analyzed in control DNA samples.
- The methylated control DNA sample seemed not to be completely methylated.
- The mean measured DNAm levels of ELOVL2, NHLRC1, and PDE4C were non-linear to the DNAm levels of control DNA samples, which could be influenced by degenerate PCR primers of those 3 markers.

### *Age correlations of markers in 10 ng of human blood-derived BS-DNA*

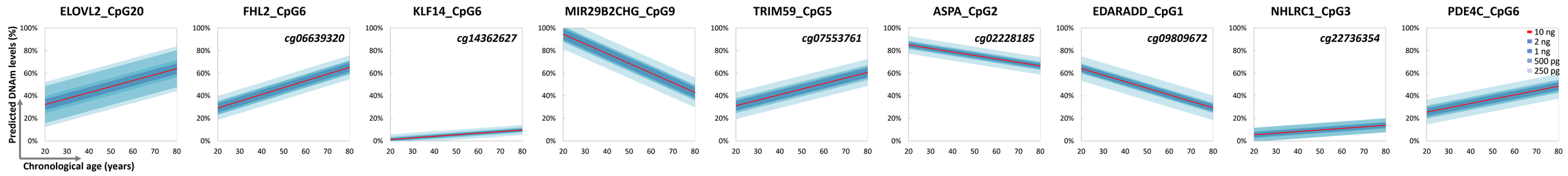



### *Increasing absolute DNAm level differences in human blood-derived BS-DNA depending on the amount of BS-DNA*




- The differences in detected DNAm levels were significantly affected by the amount of BS-DNA.
- ELOVL2 showed the largest differences in detected DNAm levels, which might be led by the low marker ratio of the marker in human blood-derived samples.
- Regarding the measured DNAm values, age-correlated CpGs in MIR29B2CHG, TRIM59, ASPA, EDARADD, and PDE4C showed around 5 percentage points of differences in the results of 1 ng BS-DNA.
- In KLF14 and NHLRC1, the detected DNAm levels can be highly affected by BS-DNA quantity since their DNAm ranges were quite low (0–20%).

### *Predicted DNAm error ranges according to the amount of BS-DNA*




## Conclusion

- The measured DNAm levels were significantly affected by the amount of BS-DNA.
- At least, more than 1 ng of BS-DNA is needed for DNAm analysis.
- Regarding the low quantity of forensic DNA samples, the robustness of age-associated markers should be considered thoroughly.

## Acknowledgement


This study was supported and funded by the Korean National Police Agency (Project Number: 0411-20220055).


## Reference


1. S.R. Hong and K.J. Shin, Frontiers in Genetics 12 (2021) 173.