DNA Methylation-based age prediction from saliva: high age predictability by combination of 7 CpG markers

Sae Rom Hong[1,2] · Sang-Eun Jung[1] · Eun Hee Lee[1] · Kyoung-Jin Shin[1,2] · Woo Ick Yang[1] · Hwan Young Lee[1]

[1]Department of Forensic Medicine, Yonsei University College of Medicine, Seoul, Korea
[2]Brain Korea 21 PLUS Project for Medical Science, Yonsei University College of Medicine, Seoul, Korea
Tel: +82-2-2228-2482, e-mail: hylee192@yuhs.ac

Introduction

DNA methylation is rising as one of the most promising age-predictive markers. Many DNA methylation-based age predictive models have been developed based on DNA methylation patterns from blood. However, a few studies have attempted to predict age from saliva, which is frequently found at crime scenes. In this study, we generated genome-wide DNA methylation profiles of saliva from 54 males and performed targeted bisulfite sequencing to identify age-associated CpGs from Saliva.

Materials and Methods

Samples

- Saliva samples from 171 males and 109 females (18 – 72 years old)
- Collected with Oragene™ DNA Self-collection Kit (DNA Genotek Inc., Ottawa, Canada)
- DNA extraction using QIAamp DNA Mini Kit (Qiagen, Hilden, Germany)
- Quantified with Quantifiler Duo Kit (Applied Biosystems, Foster city, CA, USA)

HumanMethylation450 BeadChip

- 54 saliva samples from 18 to 73 year-old males
- Infinium HumanMehtylation450 BeadChip (Illumina, San diego, CA, USA)
- To test the association between age and β-score
- Univariate linear regression analysis for each CpG site
- Selecting age-associated CpG marker candidates with criteria
  1) False dicovery rate-adjusted p-value < 0.05
  2) R^2 value > 0.65
  3) $|\ \beta\text{-score}_{MAX} - \beta\text{-score}_{min}\ | > 0.1$

Methylation SNaPshot

- 226 samples (117 males and 109 females)
- Bisulfite converted DNA by using EpiTect® Fast DNA Bisulfite Kit (Qiagen) for 500 ng of genomic DNA or Imprint® DNA Modification Kit (Sigma-Aldrich Inc., St. Louis, MO, USA) for less than 10 ng of it
- Multiplex PCR, post PCR clean up, SBE reaction and electrophoresis followed ref [1].
- Calculating methylation level by $\frac{nucleotide\ G\ intensity}{nucleotide\ G\ intensity + nucleotide\ A\ intensity}$

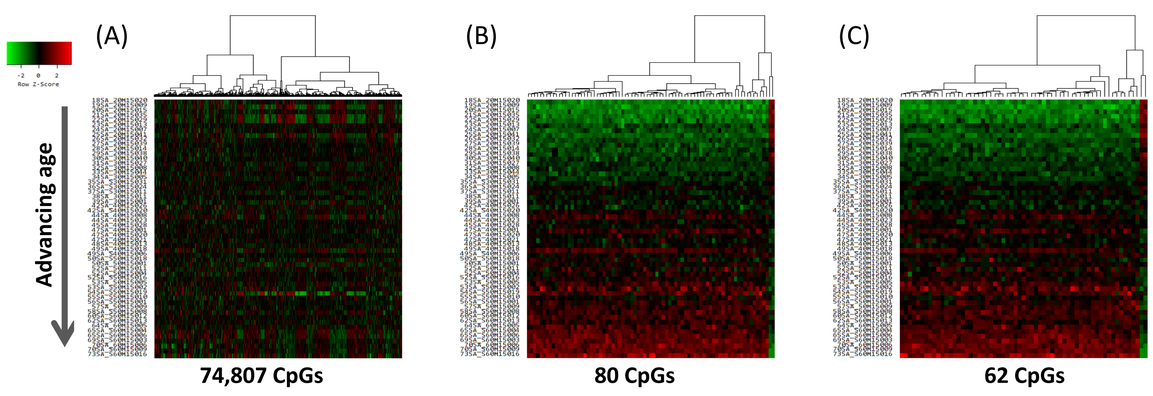
Age-predictive Model Construction

- 7 CpG sites in 226 saliva samples
- Randomly divided into 2 sets; a training set (N=113) and a testing set (N=113)
- Multivariate regression analysis to train the model using IBM SPSS Statistics for Windows, Version 23.0 (IBM Corp., Armonk, NY, USA)
- To verify the predictability of the model, the testing set was used.

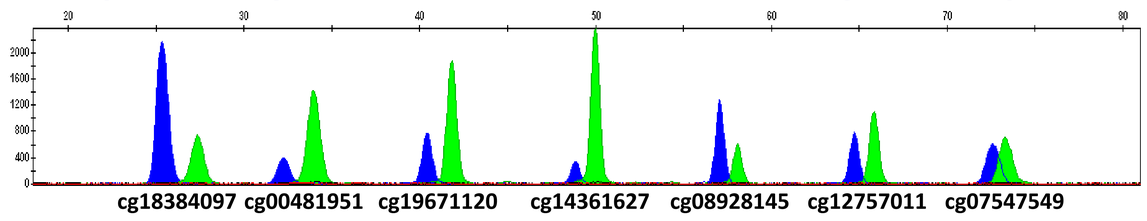
Results

Selection of age-associated CpG marker candidates

Analysis	Statistical method	No. CpGs	Cut-off	No. CpGs
Methylated level (avg.β)~ Age	Linear regression	445,791	FDR_P<0.05	74,807
			FDR_P<0.05 & R-squared>0.65	80
			FDR_P<0.05 & R-squared>0.65 & diff>=0.1	62

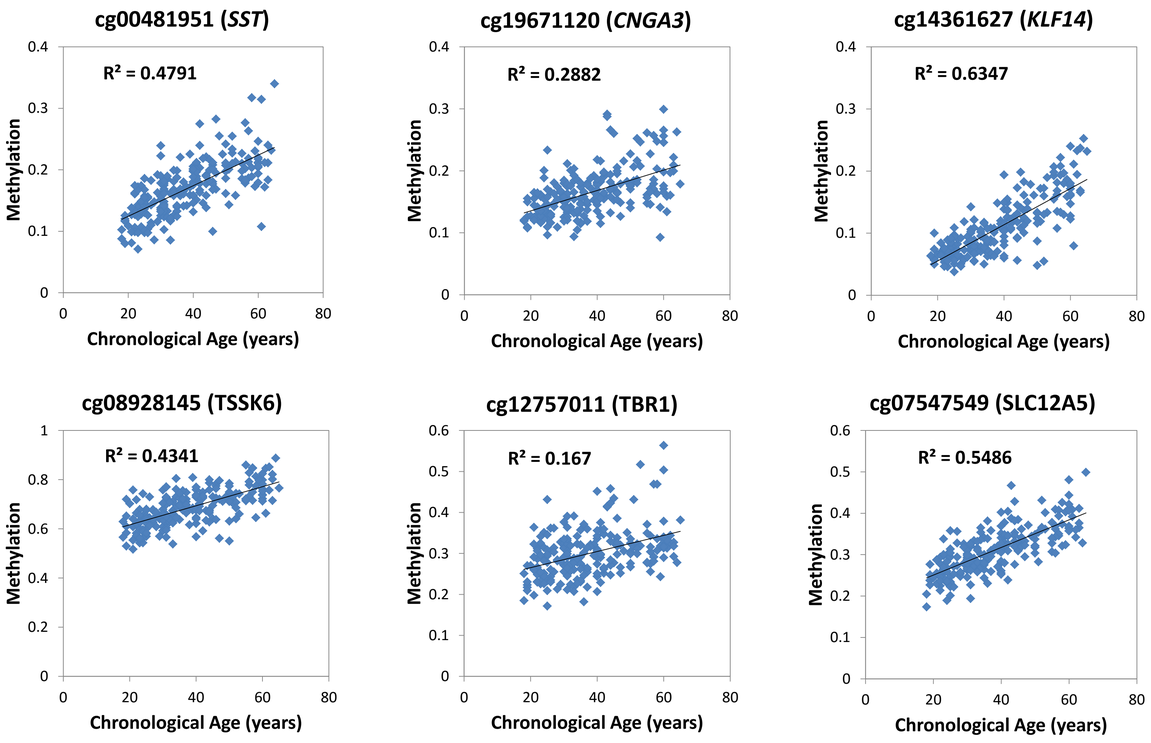


Multiplex methylation SNaPshot reaction for analysis of 7 CpGs



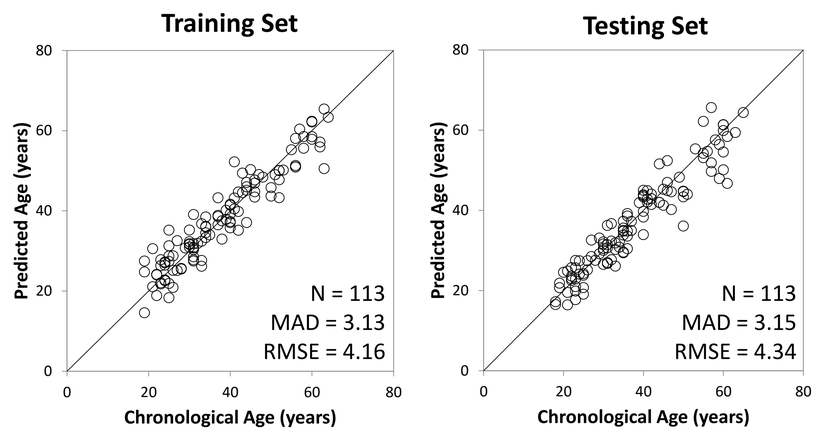
cg18384097, cg00481951, cg19671120, cg14361627, cg08928145, cg12757011, and cg07547549 are located in the *PTPN7*, *SST*, *CNGA3*, *KLF14*, *TSSK6*, *TBR1*, and *SLC12A5* genes, respectively. Green peaks represent nucleotide A as a non-methylated signal, while blue peaks represent nucleotide G as a methylation signal.

Age-associated DNA methylation changes of 6 CpG markers (N=226)



Age-predictive model for saliva

Target ID	Coefficient
(intercept)	-27.511
cg18384097	-29.088
cg00481951	9.285
cg19671120	46.992
cg14361627	86.268
cg08928145	32.211
cg12757011	58.699
cg07547549	56.384

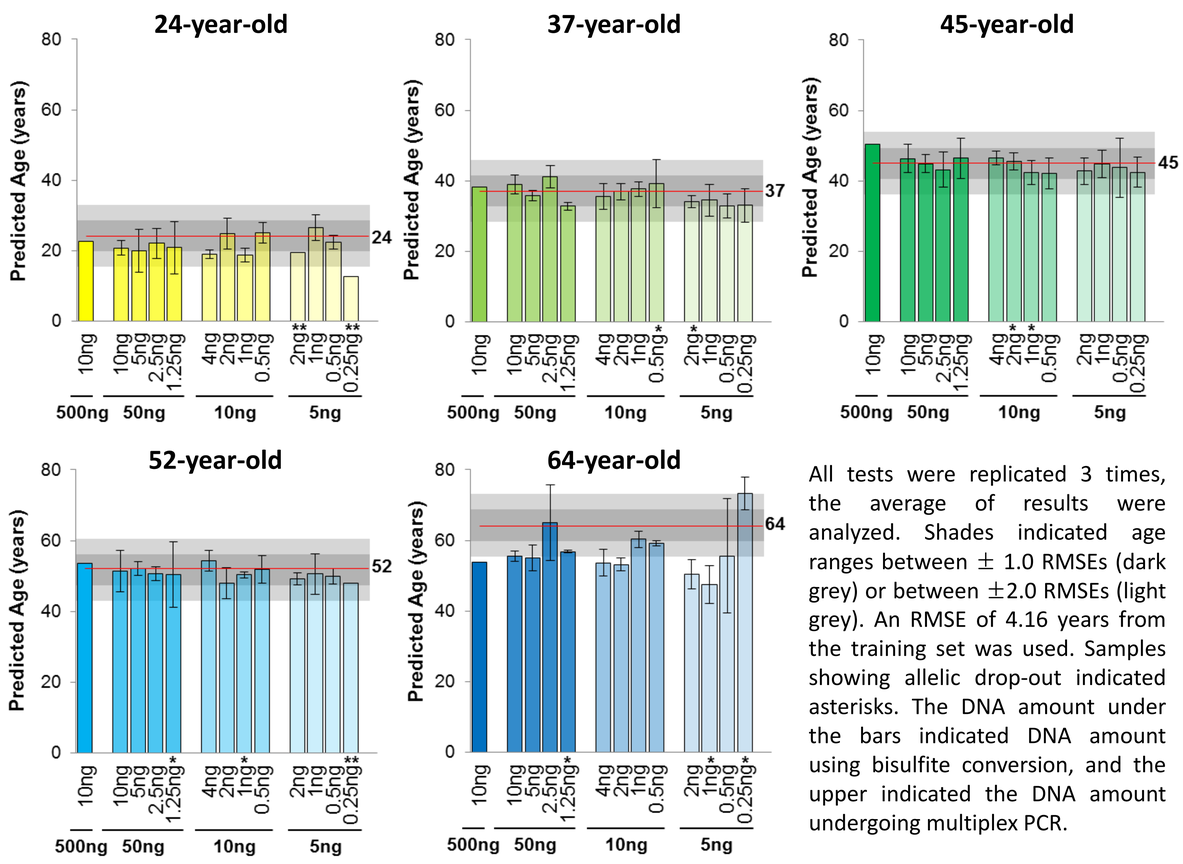


Age-predictive model performance in different age groups

Age group	Training set			Testing Set		
	No.	MAD (years)	RMSE (years)	No.	MAD (years)	RMSE (years)
20s or less	31	3.62	5.10	31	2.43	3.48
30s	32	2.39	3.69	33	2.50	3.59
40s	25	2.98	4.72	25	3.26	5.46
50s or more	25	3.63	5.48	24	4.69	7.31
Total	113	3.13	4.16	113	3.15	4.34

No., MAD, and RMSE represent the number of individuals, mean absolute deviation from chronological age, and root mean square error, respectively.

Sensitivity test for the multiplex methylation SNaPshot assay



All tests were replicated 3 times, the average of results were analyzed. Shades indicated age ranges between ± 1.0 RMSEs (dark grey) or between ±2.0 RMSEs (light grey). An RMSE of 4.16 years from the training set was used. Samples showing allelic drop-out indicated asterisks. The DNA amount under the bars indicated DNA amount using bisulfite conversion, and the upper indicated the DNA amount undergoing multiplex PCR.

Conclusion

- We analyzed DNA methylation profile of 54 saliva samples using Illumina HumanMethylation450 BeadChip array and selected 62 age-associated CpG marker candidates.
- A model composed of a cell type-specific marker (cg18384097 in *PTPN7*) and 6 age-associate markers (cg00481951 in *SST*, cg19671120 in *CGNA3*, cg14361627 in *KLF14*, cg08928145 in *TSSK6*, cg12757011 in *TBR1*, and cg07547549 in *SLC12A5*) enabled age prediction in saliva with high accuracy.
- DNA methylation profiling using the multiplex methylation SNaPshot method produced reproducible results with a small amount of DNA (4 ng of bisulfite-converted DNA). This multiplex system can be integrated into the routine forensic laboratory workflow after further validation tests with various casework samples.

Reference

[1] Hong SR, Jung SE, Lee EH, et al., Forensic Sci Int Genet. (2017) 29:118-125.

Acknowledgement

- This research was supported by fund by NRF-2014M3A9E1069992.