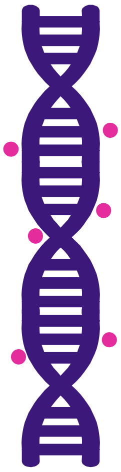
Application and Comparison of Methylation Snapshot and MPS Methods to Analyze Epigenetic Age Signatures in Saliva

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,*]

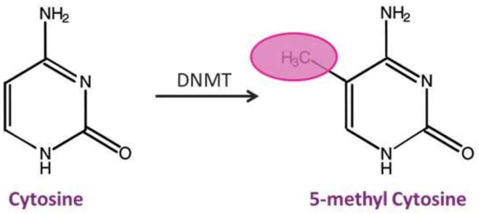
[1]Department of Forensic Medicine, Yonsei University College of Medicine, Seoul, Korea
[2]Department of Forensic Medicine and Brain Korea 21 PLUS Project for Medical Science, Yonsei University, Seoul, Korea

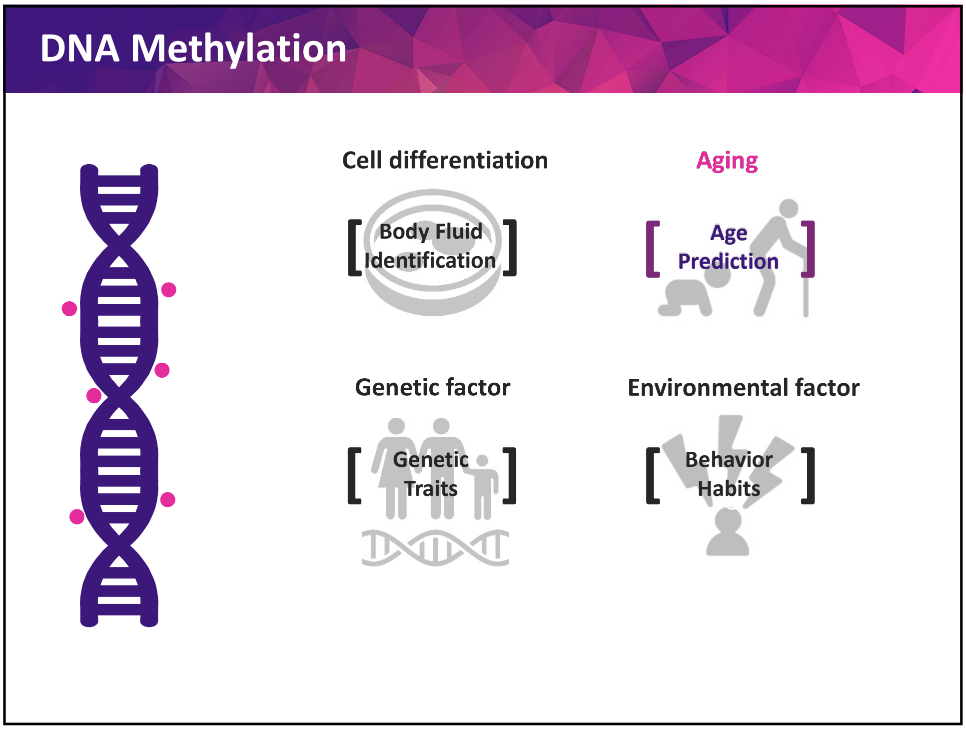
---

DNA Methylation



- Addition of a **methyl group** to **cytosine** followed by guanine
- 5'-CG-3'





DNA Methylation

Cell differentiation

[Body Fluid Identification]

Aging

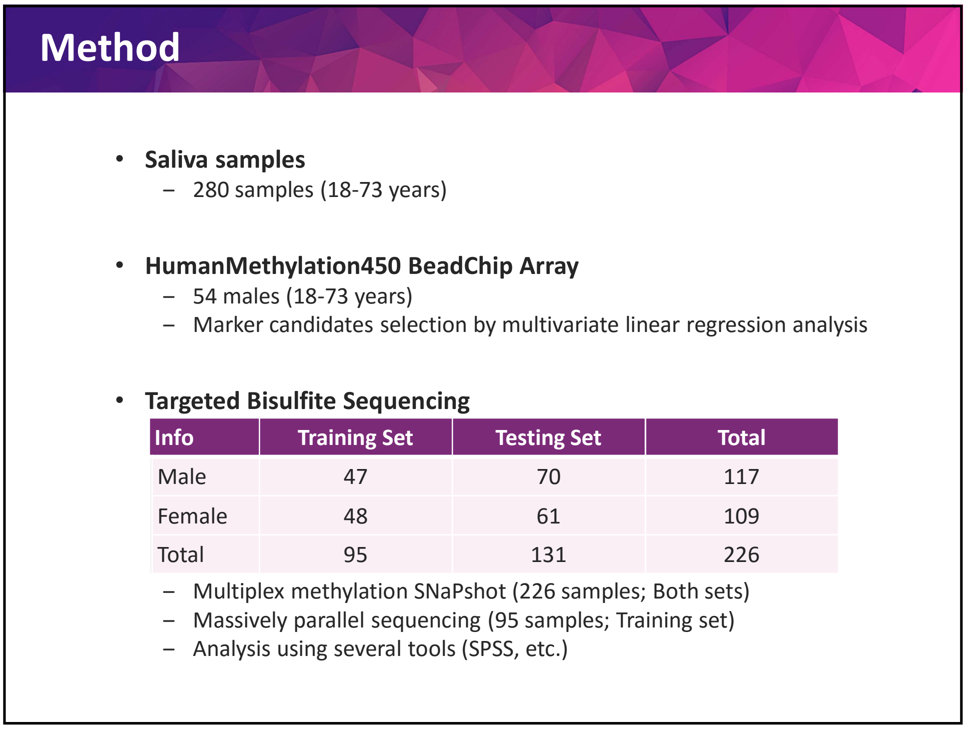
[Age Prediction]

Genetic factor

[Genetic Traits]

Environmental factor

[Behavior Habits]



Method

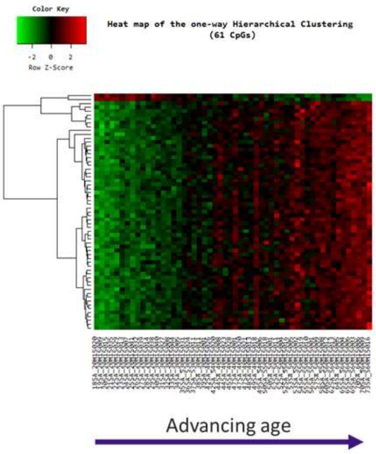
- **Saliva samples**
  - 280 samples (18-73 years)
- **HumanMethylation450 BeadChip Array**
  - 54 males (18-73 years)
  - Marker candidates selection by multivariate linear regression analysis
- **Targeted Bisulfite Sequencing**

Info	Training Set	Testing Set	Total
Male	47	70	117
Female	48	61	109
Total	95	131	226

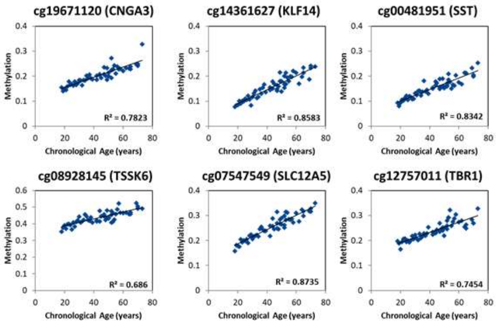
  - Multiplex methylation SNaPshot (226 samples; Both sets)
  - Massively parallel sequencing (95 samples; Training set)
  - Analysis using several tools (SPSS, etc.)

Marker Selection

HumanMethylation450 BeadChip Array



- **6 age-associated CpG candidates**
  + Cell type-specific marker (cg18384097)



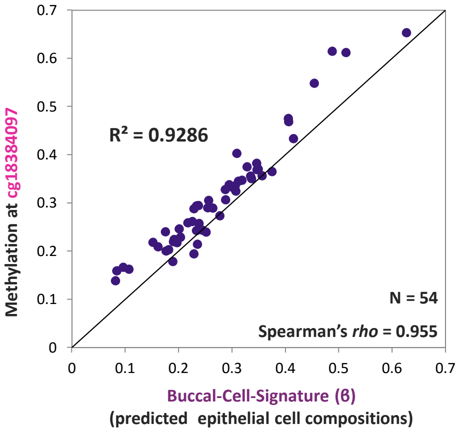
Hong *et al.* FSI Genet. (2017)

Cell Type-specific Marker

- **Buccal-Cell-Signature (β)**
  - Eipel *et al.* Aging (Albany NY). (2016)
  - cg07380416 (*CD6*)
  - cg20837735 (*SERPINB5*)
  - Percentage of buccal epithelial cells

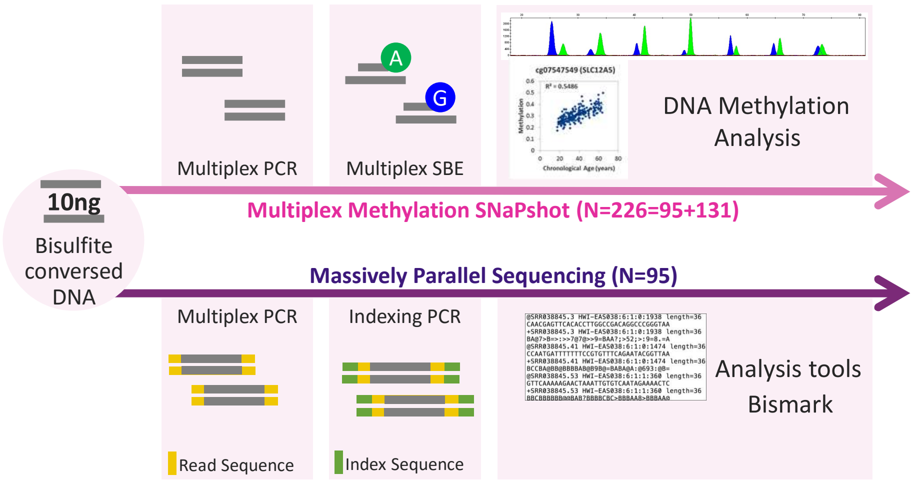
$$\beta = \frac{99.8\times\beta_{cg07380416}+1.92}{2} + \frac{-98.12\times\beta_{cg20837735}+88.54}{2}$$

- **cg18384097 (*PTPN7*)**
  - Souren *et al.* Genome Biol. (2013)
  - High in buccal epithelial cell
  - Low in blood cell
  - *PTPN7* gene
    - Protein tyrosine phosphatase (PTP)
    - Preferentially expressed in hematopoietic cells



Detail Workflow

Hong *et al*. FSI Genet. (2017)



10ng

Bisulfite conversed DNA

Multiplex PCR

Multiplex SBE

DNA Methylation Analysis

Multiplex Methylation SNaPshot (N=226=95+131)

Massively Parallel Sequencing (N=95)

Multiplex PCR

Indexing PCR

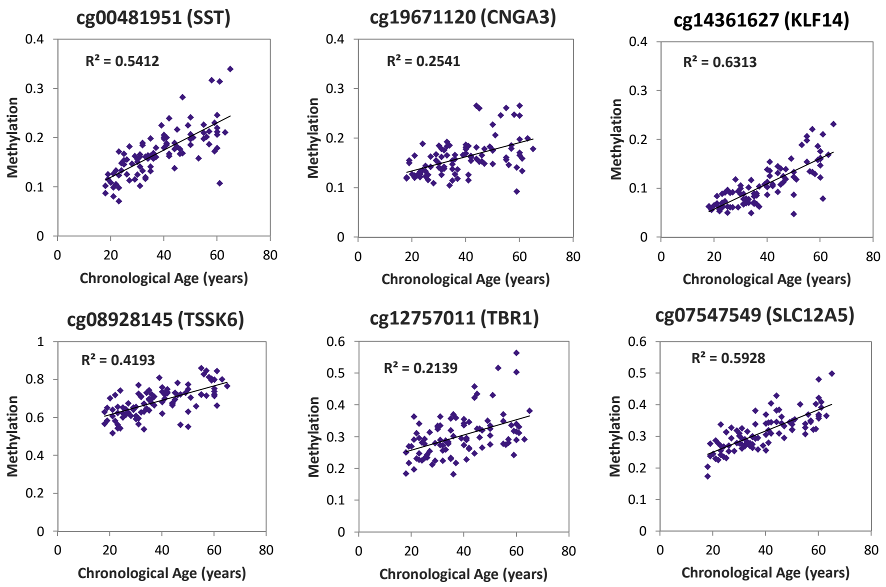
Analysis tools Bismark

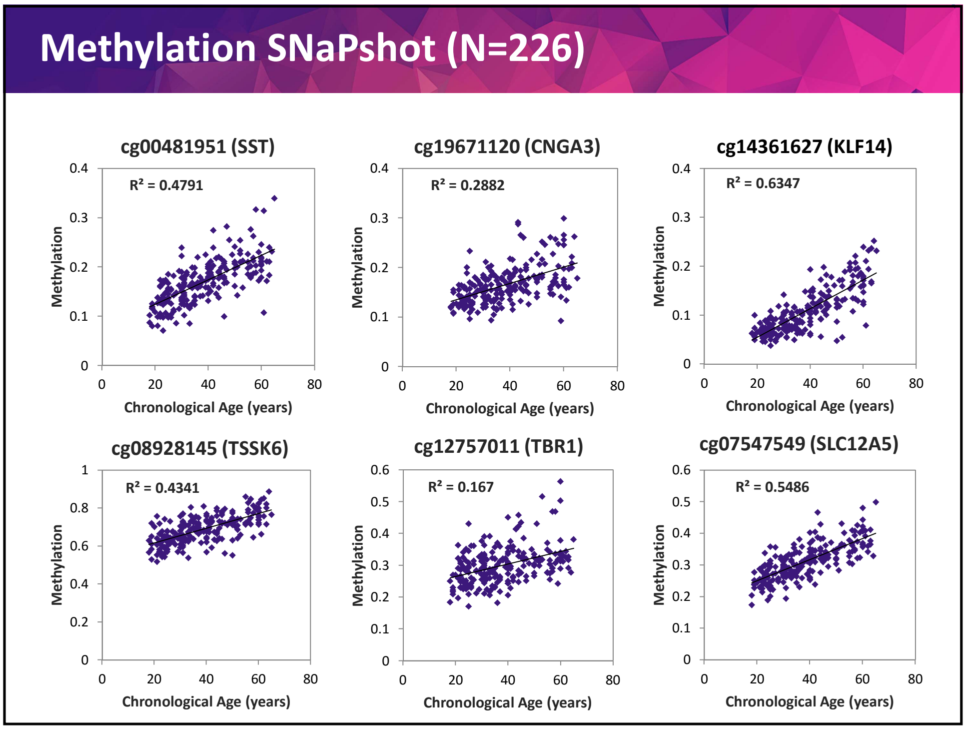
Read Sequence

Index Sequence

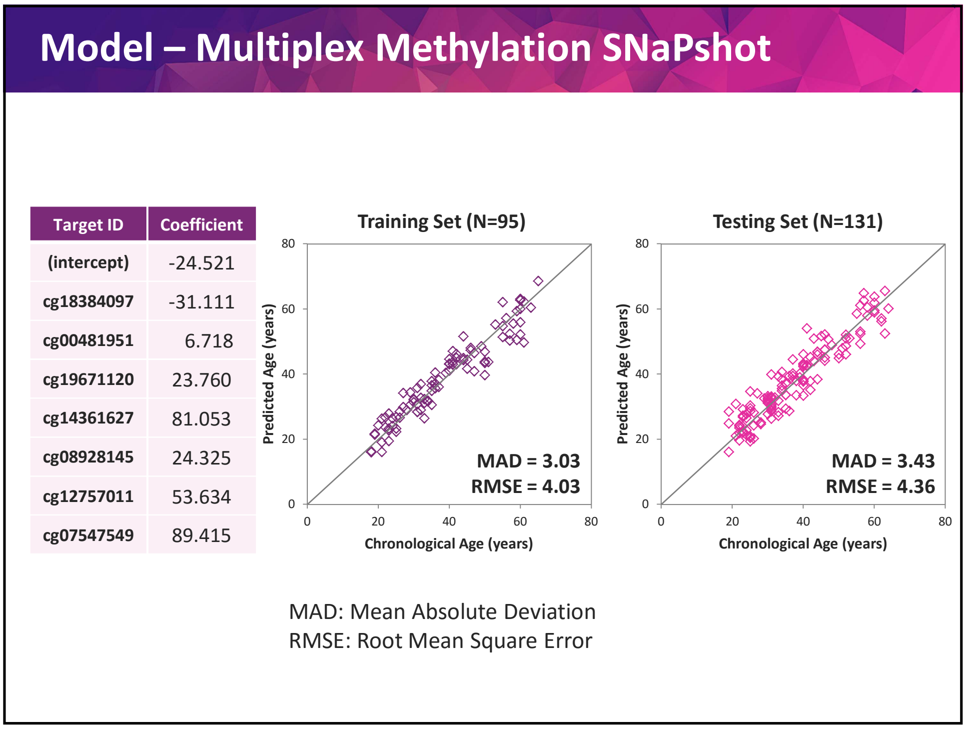
Lee *et al*. FSI Genet. (2016)

Methylation SNaPshot (N=95)





Methylation SNaPshot (N=226)



Model – Multiplex Methylation SNaPshot

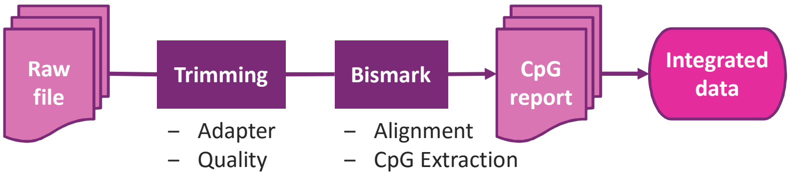
Target ID	Coefficient
(intercept)	-24.521
cg18384097	-31.111
cg00481951	6.718
cg19671120	23.760
cg14361627	81.053
cg08928145	24.325
cg12757011	53.634
cg07547549	89.415

MAD: Mean Absolute Deviation
RMSE: Root Mean Square Error

MPS Analysis

- **Platform**
  - MiSeq Reagent Kit v3 (2×300)
  - HiSeq 2000

- **Pipeline**



Trimming
- Adapter
- Quality

Bismark
- Alignment
- CpG Extraction

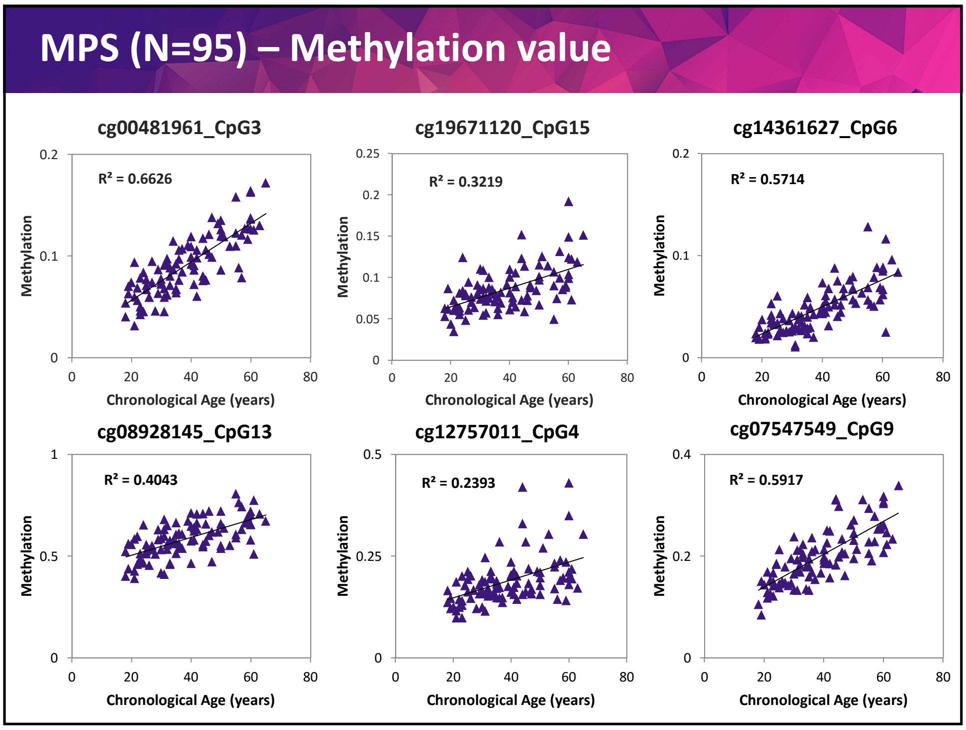
MPS (N=95) – CpG sites in amplicons

- **Pearson's R (Correlation between chonological age and methylation)**

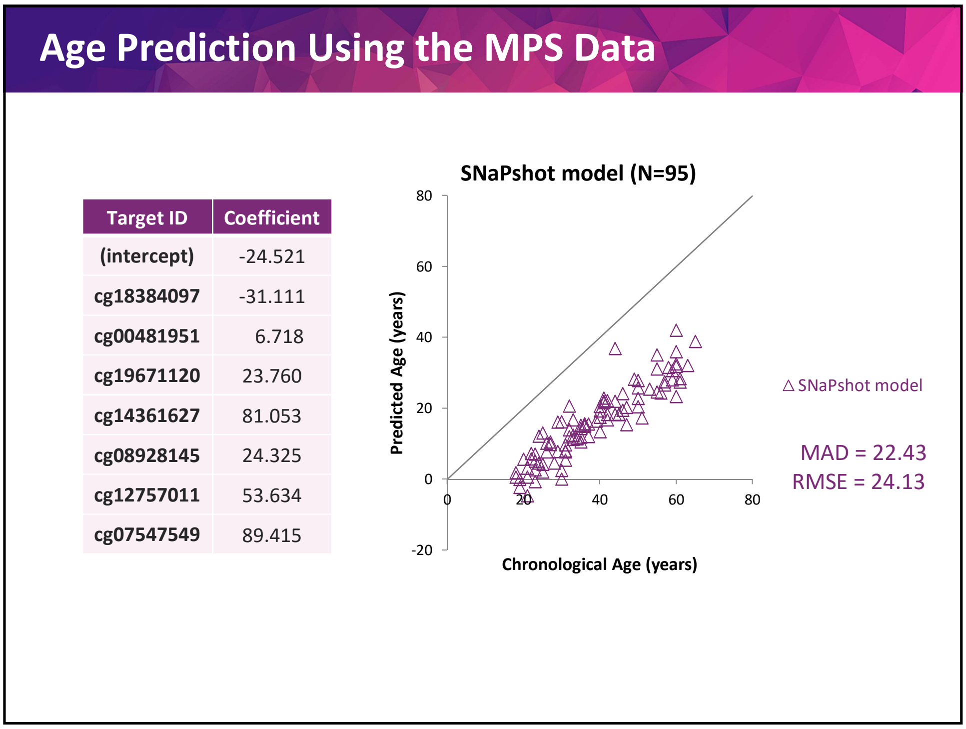
ID	CpG 1	CpG 2	CpG 3	CpG 4	CpG 5	CpG 6	CpG 7	CpG 8	CpG 9	CpG 10	CpG 11	CpG 12	CpG 13	CpG 14	CpG 15
cg18384097	-.179	-.163	-.162	-.150	-.163	**-.180**									
cg00481951	.682*	.799*	**.814***	.501*	.421*	.311*	.381*	.478*							
cg19671120	.187	.067	.104	-.033	.135	.194	.433*	.507*	.336*	.325*	.501*	.483*	.521*	.560*	**.567***
cg14361627	.261*	.492*	.556*	.631*	.650*	**.756***									
cg08928145	.596*	.584*	.662*	.649*	.637*	.637*	.616*	.637*	.641*	.629*	**.636***				
cg12757011	-.035	.229*	.319*	**.489***	-.002										
cg07547549	.321*	.130	.441*	.571*	.585*	.683*	.741*	.679*	**.769***	.399*	.285*				

■ Tageted CpG site in the methylation SNaPshot

* Statistically significant



MPS (N=95) – Methylation value

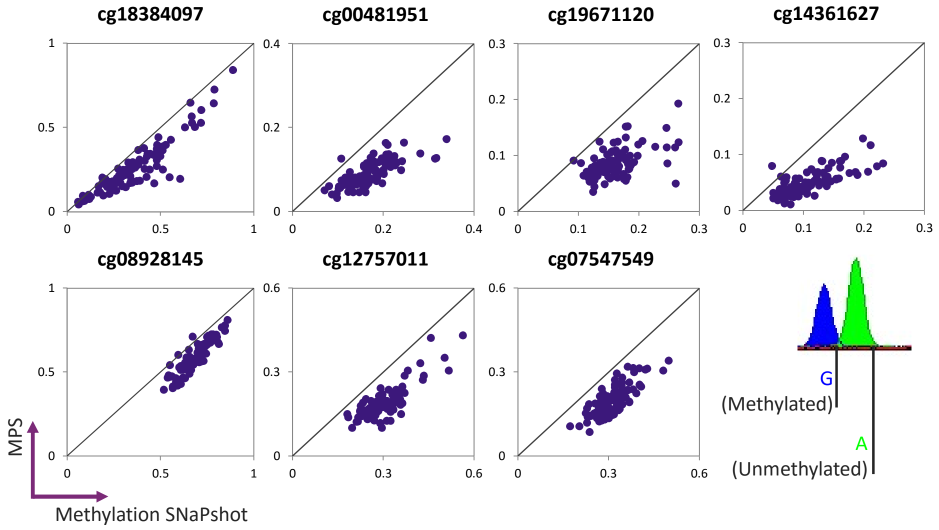


Age Prediction Using the MPS Data

Target ID	Coefficient
(intercept)	-24.521
cg18384097	-31.111
cg00481951	6.718
cg19671120	23.760
cg14361627	81.053
cg08928145	24.325
cg12757011	53.634
cg07547549	89.415

MAD = 22.43
RMSE = 24.13

Methylation SNaPshot vs MPS (N=95)

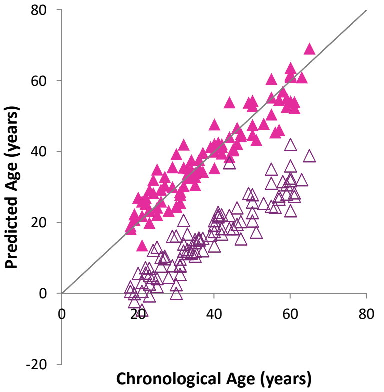


Model – MPS

Newly trained model

Target ID	Coefficient
(intercept)	-8.282
cg18384097	-20.730
cg00481951	126.188
cg19671120	77.801
cg14361627	121.858
cg08928145	20.599
cg12757011	1.820
cg07547549	78.596

MPS vs SNaPshot model(N=95)

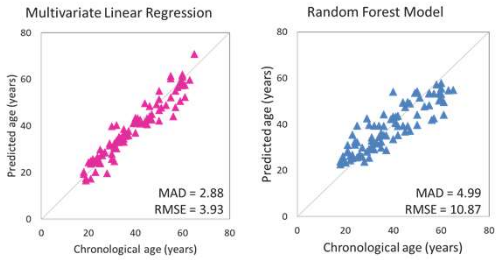


MAD = 3.59
RMSE = 4.72

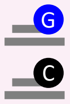
▲ MPS model
△ SNaPshot model

Further analysis

- **Analysis tool**
  - STRait razor v3.0
  - *Public available tools*
- **Various modeling**
  - Multivariate stepwise linear regression
  - Random forest modeling
  - *Other modeling*



Methylation SNaPshot vs MPS

Methylation SNaPshot	MPS
	
Multiplex	Multiplex
CE based	MPS / NGS (different platform)
Intuitive data processing	Burdensome data processing
Target CpGs only	Neighboring CpGs
Qualitative (on-off signal) Quantitative (dye intensity)	Quantitative analysis In-depth analysis

Conclusion

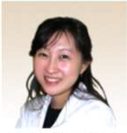
- Markers can be applied to both Multiplex methylation SNaPshot and MPS.
- The model should be altered as the platform differs.
- Models can be varied because of more information from MPS.

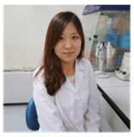
Acknowledgement

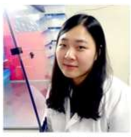
YONSEI UNIVERSITY
COLLEGE OF MEDICINE

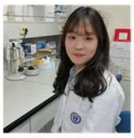
Yonsei DNA Profiling Group


Kyoung-Jin Shin


Hwan Young Lee


Sang-Eun Jung


Eun Hee Lee


Bomin Kim


Mi Hyeon Moon


SeungMin Lim

This research was supported by the Bio & Medical Technology Development Program of the National Research Foundation of Korea (NRF) funded by the Korean government (**NRF- 2014M3A9E1069992**).