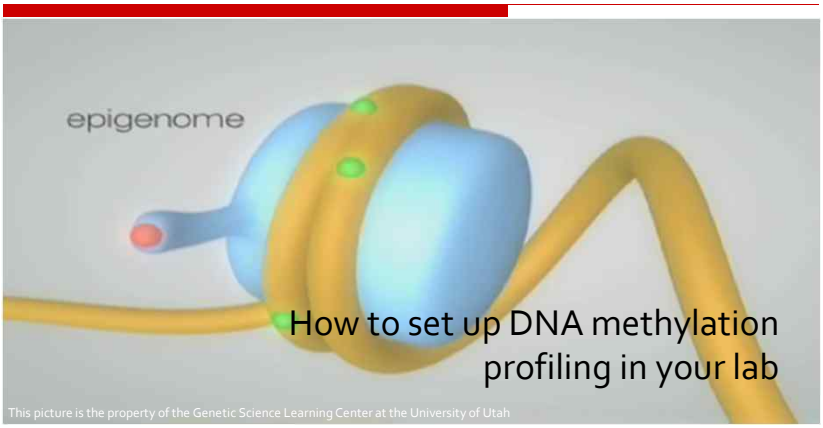
# DNA 메틸레이션의 소개와 미래

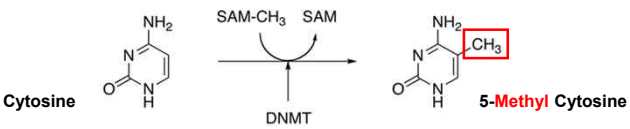



## Forensic application of DNA methylation variation

1. **Tissue-specific DNA methylation changes**
   - Body fluid identification for a clue on activity level of the evidence
2. **Age-associated DNA methylation changes**
   - Age prediction for the unknown suspect

Met

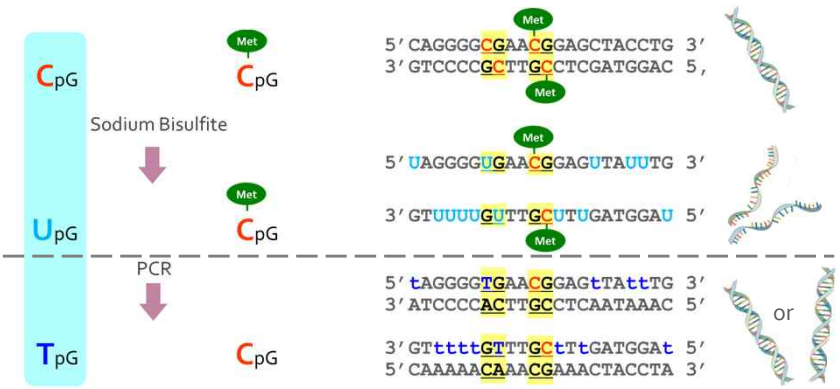
5′ CAGCTCTTCAGGGGCGAAGAGCAGGAACCGGAGCTACCTGAAGAGCG 3′



## DNA methylation and bisulfite modification

DNA methylation detection

- Bisulfite treatment converts non-methylated C to U, which changes into T by PCR

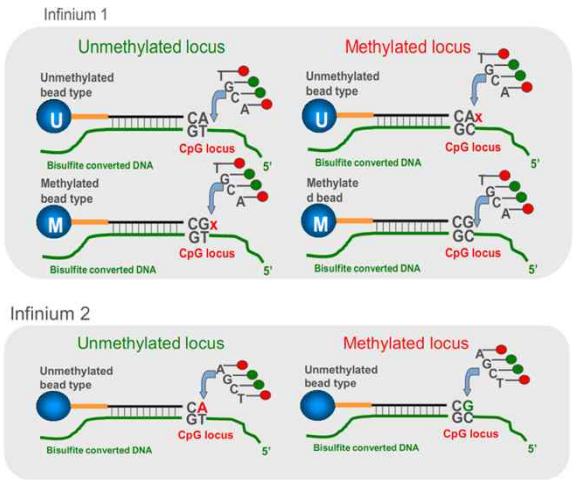



## Illumina methylation beadchip array chemistry

Array

- 27K
- 450K
- 850K: Epic array

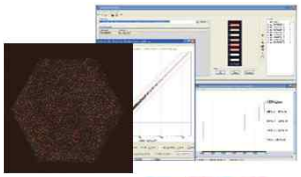
$$Beta_i = \frac{\max(y_{i,methy}, 0)}{\max(y_{i,unmethy}, 0) + \max(y_{i,methy}, 0) + \alpha}$$



## BeadChip array data analysis

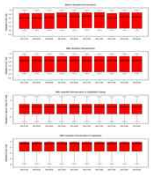
**Image analysis**

- Background Normalization
- Filtering Out CpG site with detection *p*.value > 0.05



**Data pre-processing**

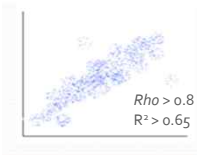
- Correlation
- Quantile Normalization
- Data Transformation
  : **Avg_beta value (0-1)**, M value



**Analysis**

- **Criteria**

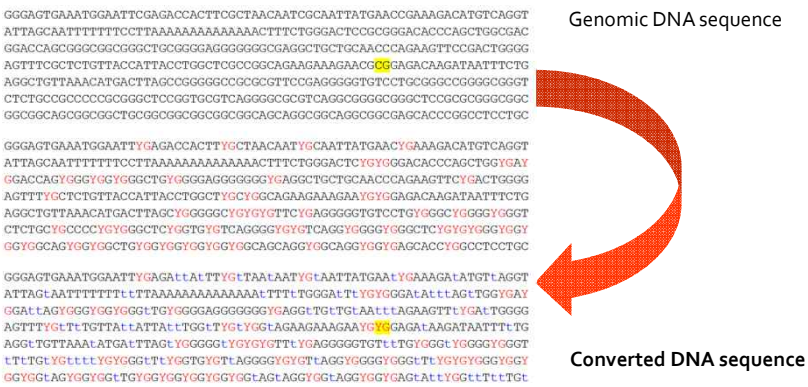
1) T-test; *p*.value < 0.05
2) |delta_mean|
3) Correlation Analysis; *rho*
4) Linear Regression; $R^2$




## DNA methylation and PCR primer design

### PCR primer design

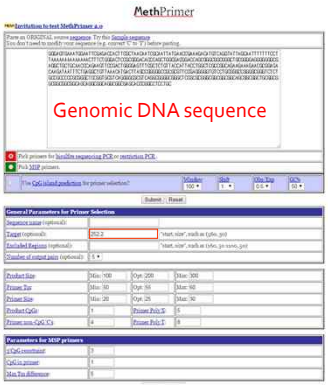
- Targets bisulfite-converted DNA



## DNA methylation and PCR primer design

PCR primer design

- **MethPrimer program (http://www.urogene.org/methprimer/)**
- PyroMark Assay Design program (Qiagen)



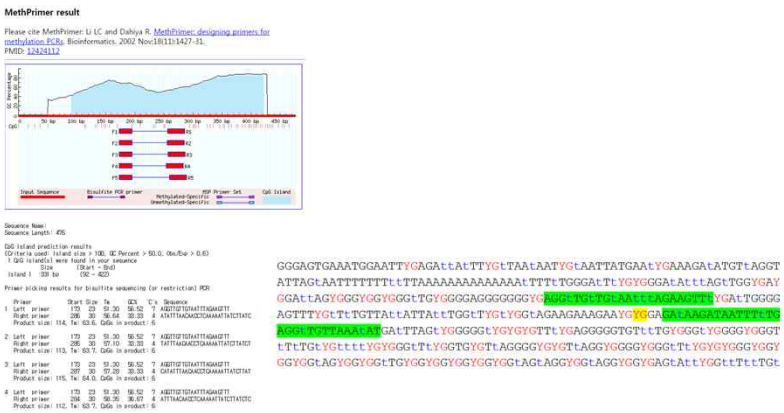

| | | | | |
|---|---|---|---|---|
| Product Size: | Min: 100 | Opt: 200 | Max: 300 | |
| Primer Tm: | Min: 50 | Opt: 55 | Max: 60 | |
| Primer Size: | Min: 20 | Opt: 25 | Max: 30 | |
| Product CpGs: | 1 | Primer Poly X: | 5 | |
| Primer non-CpG 'C's: | 4 | Primer Poly T: | 8 | |

**Primer non-CpG 'C's:** The minimum number of non-CpG 'C's in a primer. This is important for discriminating between the bisulfite-modified DNA and unmodified or incompletely modified DNA.

## DNA methylation and PCR primer design

PCR primer design

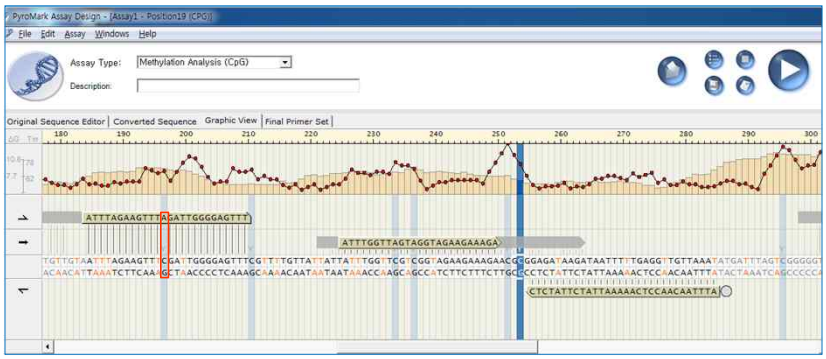
- MethPrimer program (http://www.urogene.org/methprimer/)



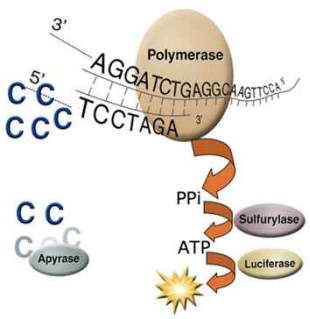
## DNA methylation and PCR primer design

PCR primer design

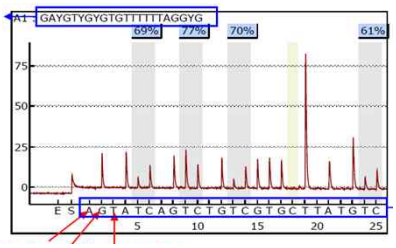
- **PyroMark Assay Design program (Qiagen)** allows mismatch at primer binding site
- YG → AG (forward primer) & CR → CC (reverse primer)




## Targeted bisulfite sequencing: pyrosequencing

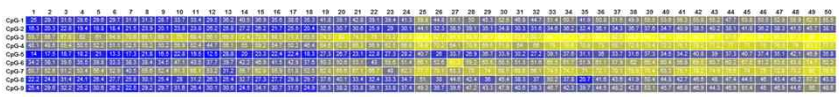



Target sequence

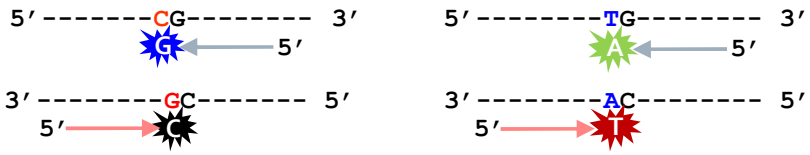



Dispensation order

$$\frac{\text{C signal intensity}}{\text{T signal + C signal}} \times 100$$



## Targeted bisulfite sequencing: methylation SNaPshot




**FORWARD**


$$\%\text{methyl} = \frac{\text{C intensity}}{\text{(C+T) intensity}} \times 100$$

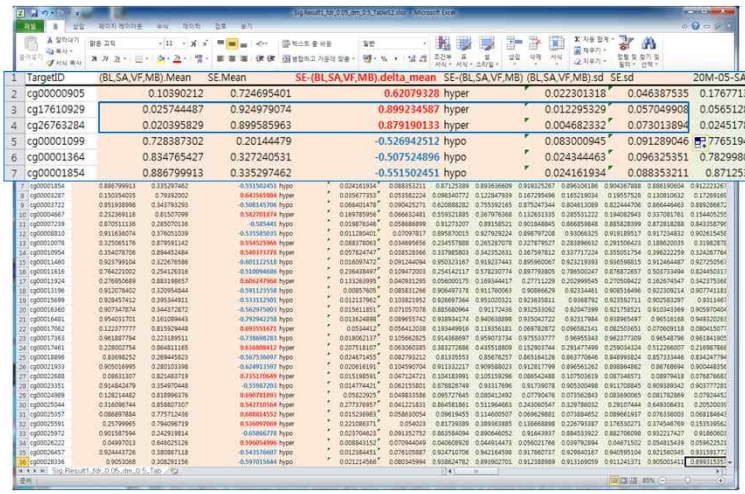
**REVERSE**


$$\%\text{methyl} = \frac{\text{G intensity}}{\text{(G+A) intensity}} \times 100$$

## Development of an assay for body fluid typing

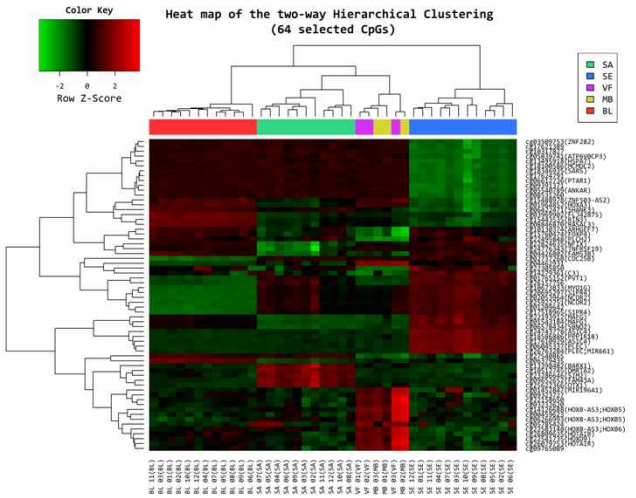
1. **Identification of body fluid-specific CpGs**
   - Genome-wide screening with Illumina's beadchip array or searching NCBI GEO database
   - 12 semen, 12 blood, 12 saliva, 6 vaginal fluid, 9 menstrual blood samples
   - Selection criteria: CpGs having methylation signal only in a target body fluid with completely non-methylation in other body fluids
   - Candidate marker validation with targeted bisulfite sequencing, such as pyrosequencing and methylation SNaPshot
2. **Design of a multiplex assay**
   - Methylation SNaPshot
   - Semen (2), blood (2), vaginal fluid(2), saliva (1), menstrual blood (2)
3. **Validation of body fluid-specific CpGs**
   - Independent samples used for marker identification
   - Set-up interpretational guidelines

## Identification of body fluid-specific CpGs



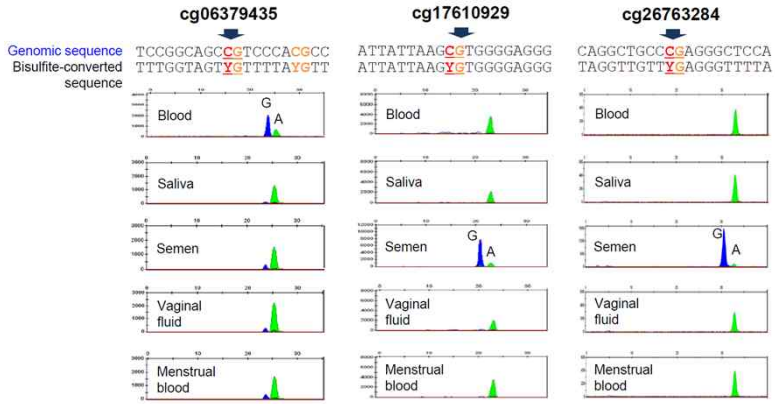
| | TargetID | (BL,SA,VF,MB).Mean | SE.Mean | SE-(BL,SA,VF,MB).delta_mean | SE-(BL,SA,VF,MB) | (BL,SA,VF,MB).sd | SE.sd | 20M-05-SA | 20M-06-SA |
|---|---|---|---|---|---|---|---|---|---|
| 2 | cg00000905 | 0.10390212 | 0.724695401 | 0.62079328 | hyper | 0.022301318 | 0.046387535 | 0.176771347 | 0.139888155 |
| 3 | cg17610929 | 0.025744487 | 0.924979074 | 0.899234587 | hyper | 0.012295329 | 0.057049908 | 0.056512882 | 0.059211692 |
| 4 | cg26763284 | 0.020395829 | 0.899585963 | 0.879190133 | hyper | 0.004682332 | 0.073013894 | 0.024517806 | 0.026326755 |
| 5 | cg00001099 | 0.728387302 | 0.20144479 | -0.526942512 | hypo | 0.083000945 | 0.091289046 | 0.776519405 | 0.754103857 |
| 6 | cg00001364 | 0.834765427 | 0.327240531 | -0.507524896 | hypo | 0.024344463 | 0.096325351 | 0.782998045 | 0.84292299 |
| 7 | cg00001854 | 0.886799913 | 0.335297462 | -0.551502451 | hypo | 0.024161934 | 0.088353211 | 0.87125389 | 0.893636609 |

## Identification of body fluid-specific CpGs

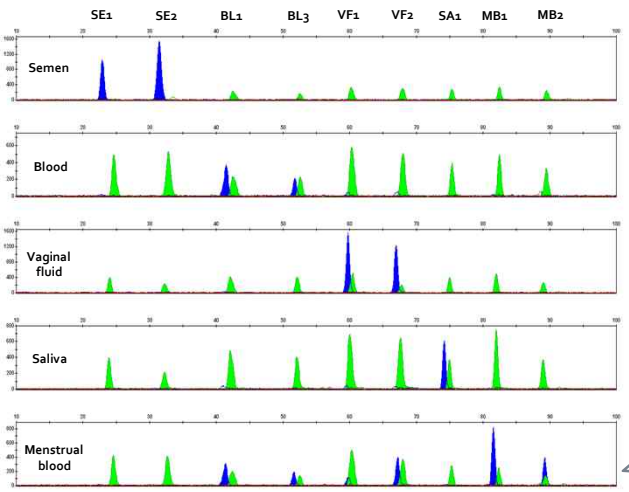


## Candidate marker validation

Targeted bisulfite sequencing using methylation SNaPshot




## A SNaPshot multiplex for body fluid identification

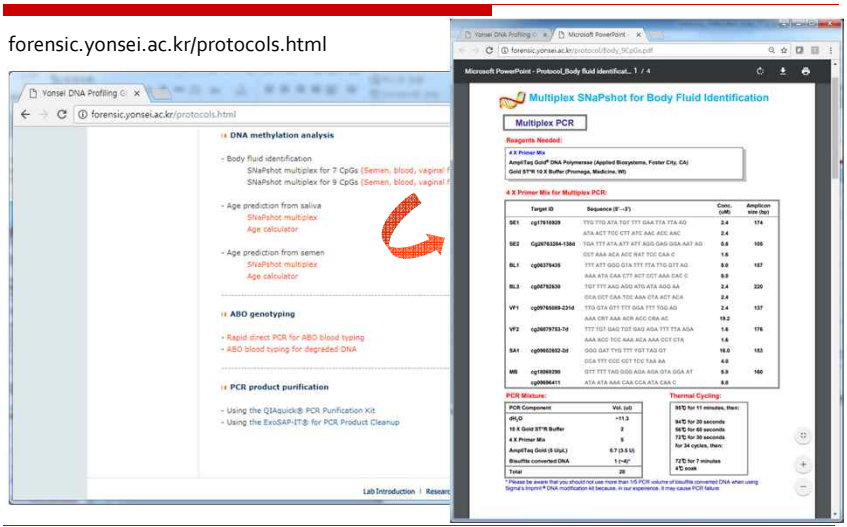



| Marker | Illumina ID | Gene |
|---|---|---|
| SE1 | cg17610929 | ACCN4 |
| SE2 | cg26763284 | PLEC |
| BL1 | cg06379435 | |
| BL3 | cg08792630 | |
| VF1 | cg09765089 | |
| VF2 | cg26079753 | HOTAIR |
| SA1 | cg09652652 | FAM43A |
| MB1 | cg18069290 | SLC26A10 |
| MB2 | cg09696411 | SLC26A10 |

Lee HY et al., Forensic Sci Int Genet (2016)

## SNaPshot multiplexes for body fluid identification

forensic.yonsei.ac.kr/protocols.html




## Introducing DNA methylation profiling in your lab

1. **Refer to previous articles e.g.:**
   - Jung SE et al. A collaborative exercise on DNA methylation-based body fluid typing. Electrophoresis. 2016;21:2759-2766.
   - Lee HY et al. DNA methylation profiling for a confirmatory test for blood, saliva, semen, vaginal fluid and menstrual blood. Forensic Sci Int Genet. 2016;24:75-82.

2. **Prepare reagents and kits such as:**
   - Bisulfite modification Kit: Imprint® DNA Modification Kit (Sigma-Aldrich Inc. Cat. No. MOD50) or EpiTect Fast Bisulfite Conversion Kit (Qiagen, Cat. No. 59104)
   - Primers (bisulfite converted DNA-specific: multiplex PCR and multiplex SBE)
   - PCR product purification: ExoSAP-IT® (USB, Cat. No. 97067-402)
   - SBE kit: SNaPshot™ Kit (Applied Biosystems, Cat. No. 4323159)
   - Post SBE: SAP or CIP
   - CE: Hi-Di Formamide, GeneScan™ 120 LIZ™ Size Standard, Matrix Standard Set DS-02, Run Module GS STR POP4 E5, POP4 (Applied Biosystems)

## DNA methylation and bisulfite modification

Each modification kit has a different capacity in the optimal or minimum DNA quantity that it can deal with

- Procedure takes 2 to 12 hours depending on the kits
- Most kits provide less than 5 hours of processing time

| Kits and manufacturers | Recommended Input | Minimum Input |
|---|---|---|
| Imprint® DNA modification kit (Sigma) | 50 pg-200 ng | > 50 pg |
| EpiTect Bisulfite kit (Qiagen) | 1 ng-1 ug | > 1 ng |
| EZ DNA Methylation™ kit (Zymo Research) | 0.5 ng-2 ug | > 500 pg |
| EZ DNA Methylation-Direct™ kit | DNA, cells, tissue | > 50 pg |



## Introducing DNA methylation profiling in your lab

### 3. Bisulfite conversion of genomic DNA

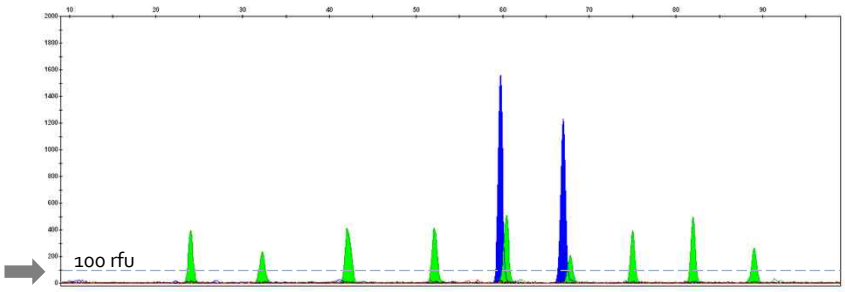
- For various body fluids: semen, blood, saliva, vaginal fluid and menstrual blood
- Extract and quantify DNA
- Select optimal input DNA amount (> 50 ng) and perform bisulfite conversion
- Bisulfite converted DNA is present as a single strand
- The converted DNA is stable for one day at room temperature, one week at 4˚C, and two to four months at –20˚C
- Recommend on storing your converted DNA below –70˚C whenever possible

### 4. Perform multiplex PCR followed by multiplex SBE

- Determine optimal amount of bisulfite converted DNA (> 10 ng)
- Perform multiplex PCR followed by multiplex SBE according to the protocol
- Analyze on CE

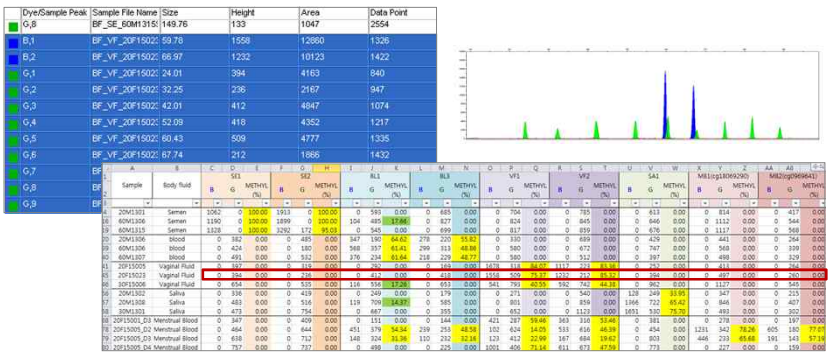
## Interpretation

- Determine detection threshold for methylation, non-methylation signal, e.g. 100 rfu
- Put peak height values into the excel sheet and calculate methylation percentages
- Determine analytical threshold for methylation percentage at each marker, e.g. 10%
- Compare with reference DNA methylation profiles to determine body fluid type




## Interpretation

- Put peak height values into the excel sheet and calculate methylation percentages
  : 100*nucleotide G /(nucleotide G+A intensities)
- Determine analytical threshold for methylation percentage at each marker, e.g. 10%
- Compare with reference DNA methylation profiles to determine body fluid type



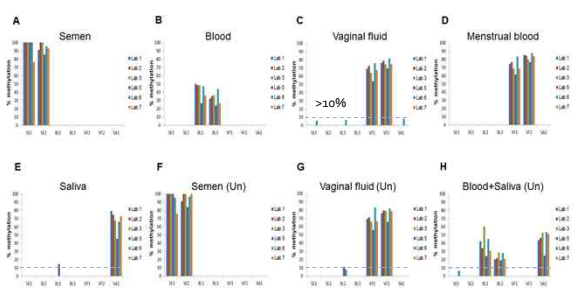
| Dye/Sample Peak | Sample File Name | Size | Height | Area | Data Point |
|---|---|---|---|---|---|
| G,8 | BF_SE_60M1315 | 149.76 | 133 | 1047 | 2554 |
| B,1 | BF_VF_20F15023 | 59.76 | 1558 | 12860 | 1326 |
| B,2 | BF_VF_20F15023 | 66.97 | 1232 | 10123 | 1422 |
| G,1 | BF_VF_20F15023 | 24.01 | 394 | 4163 | 840 |
| G,2 | BF_VF_20F15023 | 32.25 | 236 | 2167 | 947 |
| G,3 | BF_VF_20F15023 | 42.01 | 412 | 4847 | 1074 |
| G,4 | BF_VF_20F15023 | 52.09 | 418 | 4352 | 1217 |
| G,5 | BF_VF_20F15023 | 60.43 | 509 | 4777 | 1335 |
| G,6 | BF_VF_20F15023 | 67.74 | 212 | 1866 | 1432 |
| G,7 | BF | | | | |
| G,8 | BF | | | | |
| G,9 | BF | | | | |

## Interpretation

- Determine minimum threshold for positive methylation signal, e.g. 10%
- Compare with reference DNA methylation profiles to determine body fluid type

Research Article

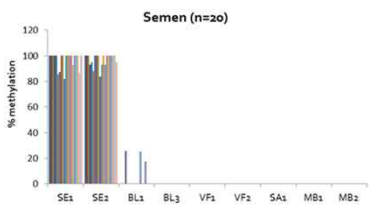
A collaborative exercise on DNA methylation based body fluid typing

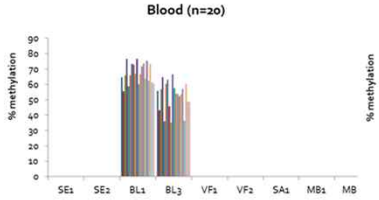


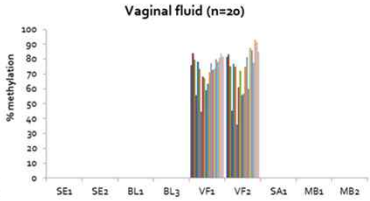

DNA methylation profiling results from 6 laboratories using bisulfite converted DNA from the same batch: a few unexpected signals with <10% methylation

## Interpretation

- Compare with reference DNA methylation profiles to determine body fluid type
  - Two markers positive for semen, blood or vaginal fluid → observed



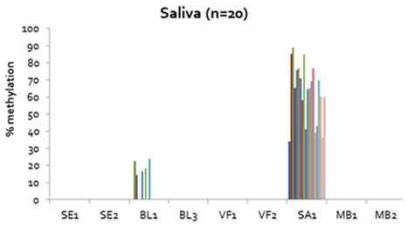



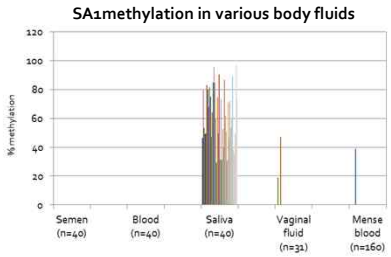




- Only one marker positive for semen, blood or vaginal fluid
  → not reliable

## Interpretation

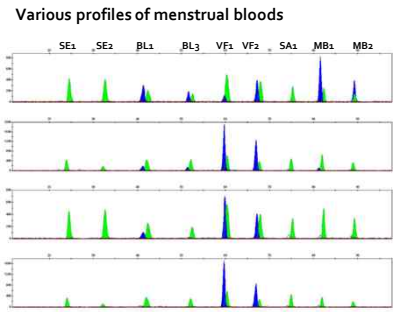
- Compare with reference DNA methylation profiles to determine body fluid type
  - Only one marker positive for saliva → observed
  - One marker positive for saliva but with additional positive signal → need more careful considerations

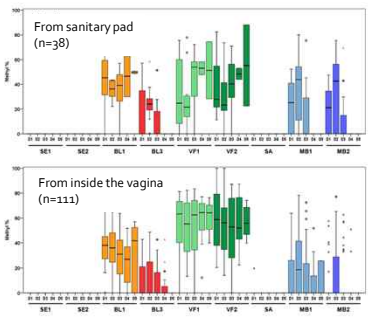



## Interpretation

- Compare with reference DNA methylation profiles to determine body fluid type
  - Menstrual blood samples show different methylation profiles depending on the menstrual cycle and the sampling methods
  - Both of the MB1 and MB2 markers positive → observed
  - One of the MB1 and MB2 markers, BL1 and BL3 markers positive → observed



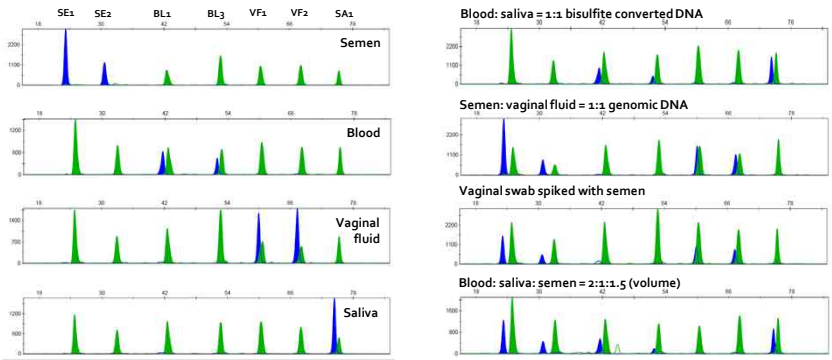



## Introducing DNA methylation profiling in your lab

5. **Analyze more samples**
   - More body fluid samples (validation and establishment of interpretational guidelines)
   - Lower amount (< 50 ng of g DNA and < 10 ng of bisulfite converted DNA)
   - Mixed samples (Joined interpretation with STR and DNA methylation profiles)
6. **Develop reporting guidelines or formats**
7. **Establish Standard Operating Procedures**
8. **Inform your clients**

## More complex samples

**e.g.** Mixed samples test in a lab of collaborative exercise (Electrophoresis, 2016)

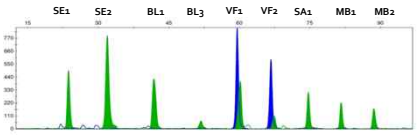


## Examples of NFS DNA methylation profiling

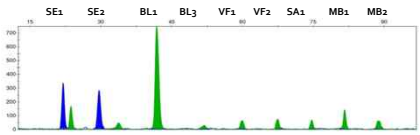
### Case 1: Rape

- Vaginal swab from the victim was positive for semen in ACP test and the A-STR profile of the perpetrator was observed
- From a penile swab, STR profile of the victim was obtained
- A stain from the belly of the victim was negative for semen and saliva, but showed STR profile of the perpetrator

**A penile swab (victim's STR profile): 0.427 ng/ul**




**A stain from the belly (perpetrator's STR profile): 0.071 ng/ul**

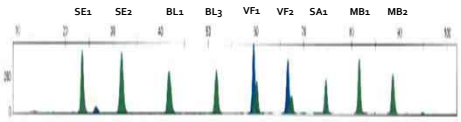



## Examples of NFS DNA methylation profiling

### Case 2: Digital penetration on the public transportation

- Vaginal swab was negative for semen and showed only STR profile of the victim
- A swab wiped the fingers of the alleged perpetrator showed STR profile of the victim

**A finger swab from the alleged perpetrator (victim's STR profile) : 0.051 ng/ul**

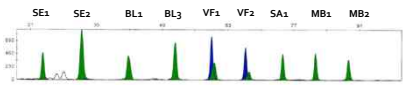


## Examples of NFS DNA methylation profiling

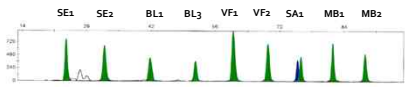
### Case 3: Murder in 2004

- A female karaoke worker was found to be dead
- A knife and an unknown stain were found near the victim
- From the test in 2004
  - The knife blade was negative for blood, but showed STR profile of the victim
  - The stain was positive for saliva, and showed unknown male's STR profile

**Knife blade (victim's STR profile): 6.7 ng/ul**




**Unknown stain: 10.7 ng/ul**



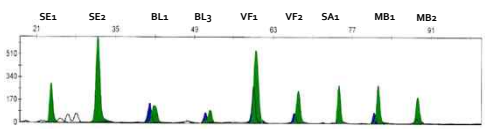

- The victim's body was damaged by cutting the left breast part and vagina with the knife
- DNA methylation profile was consistent with the scenario of the case
- DNA methylation profiling was successfully done with 13 year-old DNA samples

## Examples of NFS DNA methylation profiling

### Case 4: Murder at the scene of a fire in 2004

- A fire broke out in a room of a cafe at night
- A woman's dead body was found under a blanket with a stab wound to the neck
- A tissue paper with red stains was found at the scene
- From the test in 2004
  - The tissue paper was positive for blood and negative for semen
  - STR profile from the victim was detected on the tissue paper

**Tissue paper (victim's STR profile): 2.5 ng/ul**



## Development of an assay for age prediction

**A model-based prediction analysis typically consists of three steps**

- Discovery step
- Model-building step
- Model-validation step
- The samples in each step must be independent

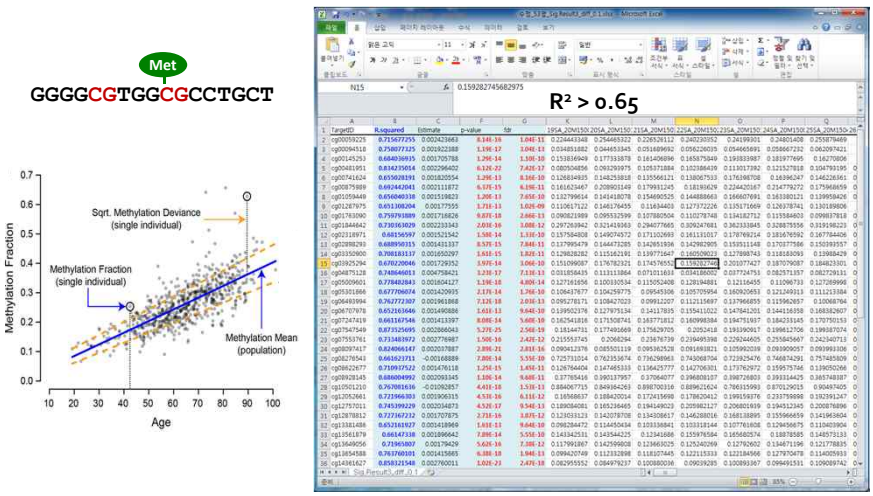
### 1. Identification of age-associated CpGs

- Regression analysis between DNA methylation and age

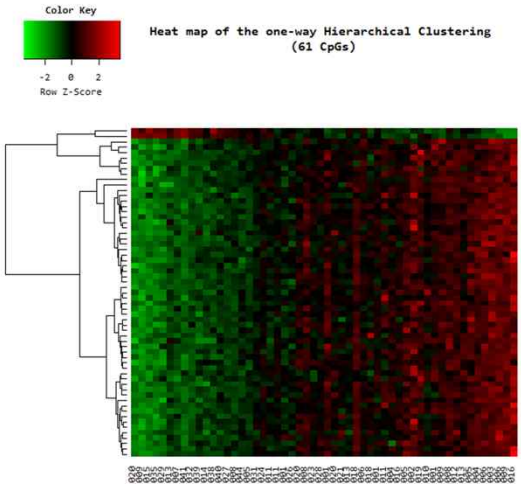
### 2. Model training and validation

- Determination of analysis platform (e.g. array, pyrosequencing, methylation SNaPshot)
- Model training: multivariate regression, random forest etc.

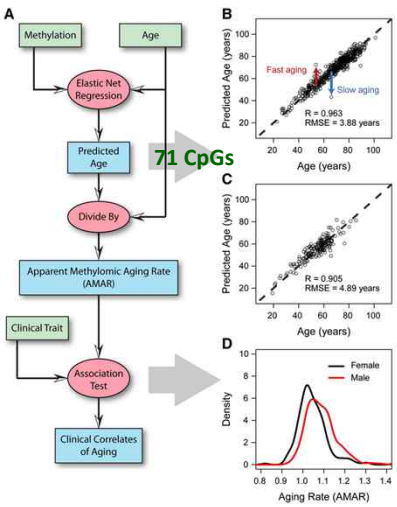
## Identification of age-associated CpGs



## Identification of age-associated CpGs




## Age prediction model construction




Regression function:

Age = a×$\underline{CpG_1}$ + b×$\underline{CpG_2}$ + c×$\underline{CpG_3}$ + …

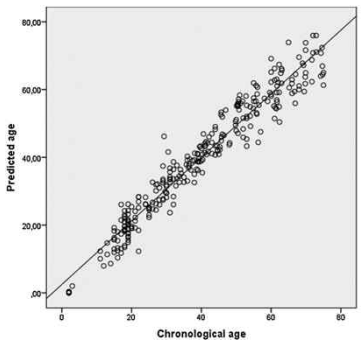
**Hannum's model: 71 CpGs**

Less than 5 year error in blood

Hannum et al. Mol Cell (2013)

## Age prediction in blood

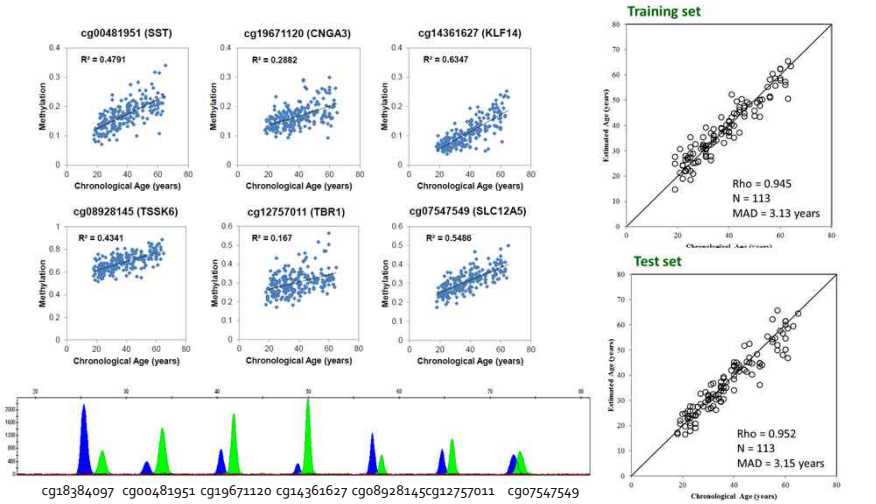
5 CpGs in the genes ELOVL2, C1orf132, TRIM59, KLF14 and FHL2




MAD (Mean Absolute Deviation) from chronological age = 3.9 years

Zbieć-Piekarska et al. Forensic Sci Int Genet (2015)

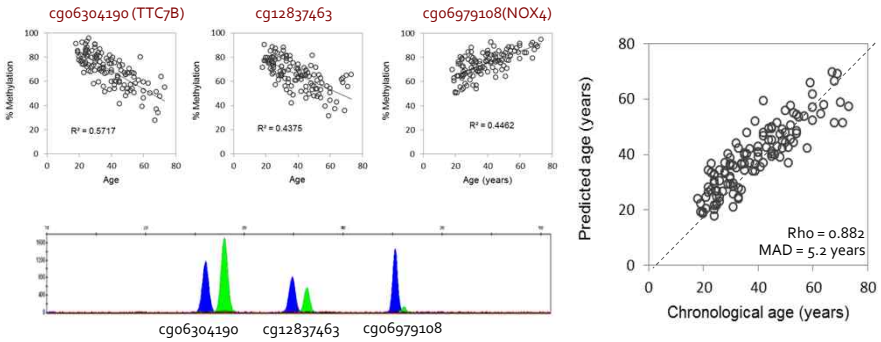
## Age prediction in saliva




Hong et al. Forensic Sci Int Genet (2017)

## Age prediction in semen

- Age correlation of the 3 CpGs and predicted versus chronological ages of 125 semen samples

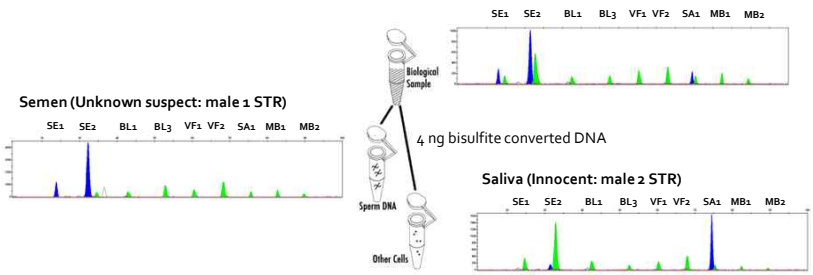



Lee et al. Forensic Sci Int Genet (2015)

## Examples of NFS DNA methylation profiling

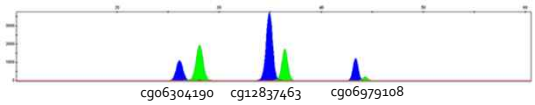
### Case 5: Indecent exposure with no suspect

- A stain preliminary positive for semen and saliva
- Two men's mixed STR profile
- Differential extraction and body fluid ID test
  - Supernatant: STR profile of a man
  - Precipitate: STR profile of another unknown man → add. age prediction



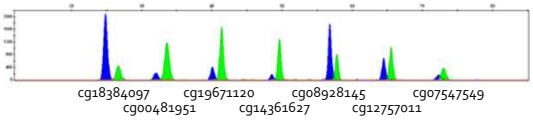
## Examples of NFS DNA methylation profiling

Semen age prediction (ppt)




61 years old

Saliva age prediction (supernatant)




30 years old

## Future forensic DNA typing

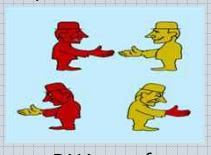
To Get More Information on Suspects and Crime Scenes

Pigmentation

Age prediction

Biogeographical Ancestry

Body fluid identification



DNA transfer

Research Networks

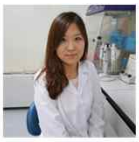
Education, Training

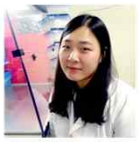
Ethics, Laws

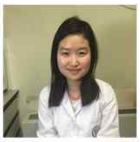
## Acknowledgement



Kyoung-Jin Shin

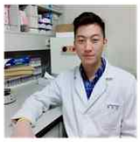

Sang-Eun Jung


Eun Hee Lee


Sae Rom Hong


Bomin Kim


Mi Hyeon Moon


SeungMin Lim

YONSEI UNIVERSITY COLLEGE OF MEDICINE

NFS NATIONAL FORENSIC SERVICE

NRF National Research Foundation of Korea 한국연구재단