Forensic Science International: Genetics 29 (2017) 250–256

Contents lists available at ScienceDirect

Forensic Science International: Genetics

journal homepage: www.elsevier.com/locate/fsig

Research paper

Independent validation of DNA-based approaches for age prediction in blood

Sohee Cho[a], Sang-Eun Jung[b], Sae Rom Hong[b,c], Eun Hee Lee[b], Ji Hyun Lee[d], Soong Deok Lee[a,d], Hwan Young Lee[b,*]

[a] Institute of Forensic Science, Seoul National University College of Medicine, Seoul, Korea
[b] Department of Forensic Medicine, Yonsei University College of medicine, Seoul, Korea
[c] Brain Korea 21 PLUS Project for Medical Science, Yonsei University, Seoul, Korea
[d] Department of Forensic Medicine, Seoul National University College of Medicine, Seoul, Korea

ARTICLE INFO

Article history:
Received 28 February 2017
Received in revised form 21 April 2017
Accepted 27 April 2017
Available online 28 April 2017

Keywords:
DNA methylation
sjTREC
Age prediction
Blood
Forensic science
Korean

ABSTRACT

Numerous molecular biomarkers have been proposed as predictors of chronological age. Among them, T-cell specific DNA rearrangement and DNA methylation markers have been introduced as forensic age predictors in blood because of their high prediction accuracy. These markers appear highly promising, but for better application to forensic casework sample analysis the proposed markers and genotyping methods must be tested further. In the current study, signal-joint T-cell receptor excision circles (sjTRECs) and DNA methylation markers located in the *ELOVL2, C1orf132, TRIM59, KLF14,* and *FHL2* genes were reanalyzed in 100 Korean blood samples to test their associations with chronological age, using the same analysis platform used in previous reports. Our study replicated the age association test for sjTREC and DNA methylation markers in the 5 genes in an independent validation set of 100 Koreans, and proved that the age predictive performance of the previous models is relatively consistent across different population groups. However, the extent of age association at certain CpG loci was not identical in the Korean and Polish populations; therefore, several age predictive models were retrained with the data obtained here. All of the 3 models retrained with DNA methylation and/or sjTREC data have a CpG site each from the *ELOVL2* and *FHL2* genes in common, and produced better prediction accuracy than previously reported models. This is attributable to the fact that the retrained model better fits the existing data and that the calculated prediction accuracy could be higher when the training data and the test data are the same. However, it is notable that the combination of different types of markers, i.e., sjTREC and DNA methylation, improved prediction accuracy in the eldest group. Our study demonstrates the usefulness of the proposed markers and the genotyping method in an independent dataset, and suggests the possibility of combining different types of DNA markers to improve prediction accuracy.

1. Introduction

Age is one of the most valuable features for forensic investigations, as it can provide clues to the appearance of an unknown person and narrow down the number of potential suspects. Diverse biological changes that occur with respect to age in human bodies have been proposed as possible age-predictive markers, such as telomere shortening, and accumulation of mitochondrial DNA mutations or deletions [1–3]. However, these markers demonstrate low prediction accuracy, and are associated with certain technical problems that make them difficult to use in forensic laboratories.

In recent years, signal-joint T-cell receptor excision circles (sjTRECs) have been introduced in the forensic field as a useful marker for age prediction from blood samples [4–7]. The quantitative amount of this DNA molecule in human bodies changes depending on an individual's immunological status, and decreases with aging [8,9]. The prediction accuracy of this test is relatively high, with an error of 7 to 12 years [4–7,10]. Considering that aging is a complex process that is influenced by many factors, including genetic, as well as developmental and physical states [3,11,12], the accuracy of the assay might be further improved by an approach that uses multiple predictors derived from independent sources [10].

* Corresponding author at: Department of Forensic Medicine, Yonsei University College of Medicine, 50-1 Yonsei-ro, Seodaemun-gu, Seoul 03722, South Korea.
E-mail addresses: hylee192@yuhs.ac, hylee192@gmail.com (H.Y. Lee).

http://dx.doi.org/10.1016/j.fsigen.2017.04.020
1872-4973/

Meanwhile, many researchers in the fields of forensics and genetics have reported that analysis of DNA methylation is the most reliable method of predicting age, with an accuracy of a mean absolute deviation (MAD) from chronological age of less than 5 years [13–18]. Unlike the method that detects a decreasing amount of sjTRECs in blood, DNA methylation-based age prediction has an inherent advantage, in that it is applicable across a broad spectrum of tissues, including forensically-relevant body fluids, such as blood, saliva, and semen [19–22]. To date, several age-associated CpGs have been identified through genome-wide methylation analysis, and have also been validated using several different analysis platforms, e.g., pyrosequencing, EpiTyper, and methylation SNaPshot [16,22,23]. Multiple positions in the *ELOVL2* gene have often been suggested to be the most powerful known age predictors in blood despite slight differences in the strength of the age association among different CpG sites depending on the tested sample set [18,24–26]. In the forensic field, Zbieć-Piekarska et al. [16] and Park et al. [18] proposed different age prediction models based on the pyrosequencing results of several CpGs in the *ELOVL2* gene and additional CpGs in other genes, though both models exhibited very high prediction accuracy (MAD from chronological age of less than 4 years).

Most studies have focused on the identification of a new set of markers in a specific population group, but researchers have rarely replicated previous results or considered technical problems that may have occurred related to experimental design. For better application of developed approaches to forensic casework sample analysis, it is important to further test the proposed markers and genotyping methods. In the present study, we re-analyzed the sjTREC and DNA methylation markers proposed by Zbieć-Piekarska et al. [16] (CpGs in the *ELOVL2, C1orf132, TRIM59, KLF14*, and *FHL2* genes) in 100 Korean blood samples using the same methods as described in previous publications, but with different primers for target amplification. Then, based on the quantification results for sjTREC and DNA methylation in the 5 genes, we calculated the predicted age using previously reported or newly generated models, and validated the models' predictive performance with Spearman's *rho* value, and the MAD from chronological age.

2. Materials and methods

2.1. Sample collection and DNA extraction

Blood samples were collected from 100 healthy Korean volunteers of 20 to 74 years of age (50 males and 50 females) following the procedures approved by the Institutional Review Board of Seoul National University Hospital Biomedical Research Institute and Yonsei University Severance Hospital. Ninety eight out of 100 donors reside in Seoul (n = 80) or the adjacent Kyounggi province (n = 18), and 2 live in either Pusan or the Chungbuk province. Peripheral blood samples were collected in purple-topped EDTA vacutainers (BD Bioscience, San Jose, CA, USA). DNA was extracted from 300 μl of whole blood using a Maxwell® 16 Blood DNA Purification Kit (Promega, Madison, WI, USA) and a Maxwell® 16 instrument (Promega), following the manufacturer's instructions. The extracted DNA was quantified using the Quantifiler® Human DNA Quantification Kit (Thermo Fisher Scientific, Waltham, MA, USA) and stored at −20°C until use.

2.2. Real-time PCR for sjTREC quantification

The amount of sjTRECs in each sample was determined as described in a previous report [6]. A plasmid encoding the sjTREC sequence was serially diluted to 10^{10}–10^{3} copies to generate a standard curve, while 0.1 pg of the plasmid corresponding to 22,727 copies [6] was run in parallel with extracted DNA samples in each reaction as a control. Real-time PCR was performed using QuantStudio 5 (Thermo Fisher Scientific) and the results were analyzed with QuantStudio™ Design & Analysis Software 1.3.1 from the same manufacturer. Each reaction contained 50 ng of genomic DNA (gDNA), 0.8 μM of each primer, 0.25 μM of TaqMan® probe, and 1 × TaqMan® Universal PCR Master Mix II (Thermo Fisher Scientific). The sequences were 5′-TGCTGACACCTCTGGTTTTTGTAA-3′ (forward primer), 5′-GTGCCAGCTGCAGGGTTTAG-3′ (reverse primer), and 5′-FAM-CACGGTGATGCATAGGCACCTGC-TAMRA-3′ (TaqMan probe). The thermal cycling conditions were first 50°C for 2 min and then 95°C for 10 min, followed by 45 cycles at 95°C for 15 s and 60°C for 1 min. The amount of sjTREC was calculated from the C_t value and the standard curve regression equation. All samples were run twice, and the average from these results was used for statistical analysis. To validate the prediction accuracy of the previously reported model [6], predicted ages were calculated using the C_t value obtained from the present study and the age-predictive function of the previous report.

2.3. Bisulfite modification and pyrosequencing for DNA methylation analysis

Five hundred nanograms of extracted gDNA from blood was prepared for bisulfite modification, using the EZ DNA Methylation-Lightning™ kit (Zymo Research, Irvine, CA, USA), as per the manufacturer's instructions. A gDNA sample applied to CT conversion reagent in a total volume of 20 μl was incubated at 98°C for 8 min and 54°C for 60 min, stored at 4°C for up to 20 h and then purified using the reagents obtained from the same kit. The converted gDNA samples were finally stored at −20°C, until use. Methylation analysis of the targeted genes was performed by pyrosequencing. Each primer was designed using PyroMark® Assay Design Software v2.0 (Qiagen, Hilden, Germany), and is listed in Supplementary Table S1. The PCR reaction was carried out in a volume of 20 μl containing 2 μl of converted gDNA (1/25 of the eluate aliquot corresponding to approximately 20 ng), 1 × PCR premixture (Enzynomics, Daejeon, Korea), and 0.5 μM of each of the primers, according to the following amplification conditions: denatured at 95°C for 10 min, followed by 45 cycles at 95°C for 30 s, 60–62°C for 30 s, and 72°C for 30 s, and a final extension at 72°C for 5 min. The template prepared from biotinylated PCR product was sequenced with a PyroMark ID system with the Pyro Gold reagents kit (Qiagen), according to the manufacturer's instructions. The generated pyrograms were automatically analyzed using PyroMark analysis software.

2.4. Statistical analysis

Statistical analysis was performed using IBM SPSS Statistics 20 (IBM, Armonk, NY, USA), and plots were generated using SigmaPlot 12.0 software (Systat Software Inc., San Jose, CA, USA). The age related correlations of sjTREC copy numbers and methylation level at 32 CpG sites were assessed using Spearman's correlation coefficient. To generate age prediction models, multivariate linear regression analysis was performed for select CpG positions, all tested CpG positions, and/or C_t values for quantification of sjTRECs. The MAD from the chronological age was calculated for all models built from each analysis.

3. Results

3.1. Validation of an age predictive model based on the quantification of sjTRECs

The sjTREC level ($\log_{10}$[sjTREC copies per μg DNA]) quantified from the 100 Korean volunteers 20 to 74 years of age was used in