Forensic Science International: Genetics 33 (2018) 1–9

Contents lists available at ScienceDirect

Forensic Science International: Genetics

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

Research paper

DNA methylation markers in combination with skeletal and dental ages to improve age estimation in children

Lei Shi[a,c,e], Fan Jiang[b,c], Fengxiu Ouyang[a,c], Jun Zhang[a,c], Zhimin Wang[d], Xiaoming Shen[a,b,c,*]

[a] XinHua Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China
[b] Shanghai Children's Medical Center, Shanghai Jiao Tong University School of Medicine, Shanghai, China
[c] MOE-Shanghai Key Laboratory of Children's Environmental Health, Shanghai, China
[d] Department of Genetics, Shanghai-MOST Key Laboratory of Health and Disease Genomics, Chinese National Human Genome Center and Shanghai Academy of Science & Technology, Shanghai, China
[e] Department of Pediatric Dentistry, Ninth People's Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China

ARTICLE INFO

Keywords:
Age estimation
Combined methods
DNA methylation markers
Dental development
Hand-wrist bones

ABSTRACT

Age estimation is critical in forensic science, in competitive sports and games and in other age-related fields, but the current methods are suboptimal. The combination of age-associated DNA methylation markers with skeletal age (SA) and dental age (DA) may improve the accuracy and precision of age estimation, but no study has examined this topic. In the current study, we measured SA (GP, TW3-RUS, and TW3-Carpal methods) and DA (Demirjian and Willems methods) by X-ray examination in 124 Chinese children (78 boys and 46 girls) aged 6–15 years. To identify age-associated CpG sites, we analyzed methylome-wide DNA methylation profiling by using the Illumina HumanMethylation450 BeadChip system in 48 randomly selected children. Five CpG sites were identified as associated with chronologic age (CA), with an absolute value of Pearson's correlation coefficient (r) > 0.5 ($p < 0.01$) and a false discovery rate < 0.01. The validation of age-associated CpG sites was performed using droplet digital PCR techniques in all 124 children. After validation, four CpG sites for boys and five CpG sites for girls were further adopted to build the age estimation model with SA and DA using multivariate linear stepwise regressions. These CpG sites were located at 4 known genes: *DDO, PRPH2, DHX8,* and *ITGA2B* and at one unknown gene with the Illumina ID number of 22398226. The accuracy of age estimation methods was compared according to the mean absolute error (MAE) and root mean square error (RMSE). The best single measure for SA was the TW3-RUS method (MAE = 0.69 years, RMSE = 0.95 years) in boys, and the GP method (MAE = 0.74 years, RMSE = 0.94 years) in girls. For DA, the Willems method was the best single measure for both boys (MAE = 0.63 years, RMSE = 0.78 years) and girls (MAE = 0.54 years, RMSE = 0.68 years). The models that incorporated SA and DA with the methylation levels of age-associated CpG sites provided the highest accuracy of age estimation in both boys (MAE = 0.47 years, $R^2 = 0.886$) and girls (MAE = 0.33 years, $R^2 = 0.941$). Cross validation of the results confirmed the reliability and validity of the models. In conclusion, age-associated DNA methylation markers in combination with SA and DA greatly improve the accuracy of age estimation in Chinese children. This method may be applied in forensic science, in competitive sports and games and in other age-related fields.

1. Introduction

Chronological age (CA) is of great relevance with regard to social justice, such as the determination of social benefits or criminal penalties, school admission, and eligibility for sport competitions [1]. Validation of age is indispensable under these circumstances. An accurate estimation of age becomes pivotal when self-reported age is questionable. For example, in forensic practice, age estimation is always performed in evaluating the capacity for civil conduct [2], criminal responsibility [3], immigration and adoption [4,5]. For sports, CA is often applied to define eligibility and assign competing groups in national and international games [6]. An ideal method of age estimation is to

Abbreviations: SA, skeletal age; DA, dental age; MAE, mean absolute error; RMSE, root mean standard error; CA, chronological age; GP, Greulich and Pyle; TW3, Tanner and Whitehouse version 3; ddPCR, digital droplet PCR; FDR, false discovery rate; EA, estimated age

* Corresponding author at: MOE-Shanghai Key Laboratory of Children's Environmental Health, Xinhua Hospital, Shanghai Jiao Tong University School of Medicine, 1665 Kong Jiang Road, Shanghai, 200092, China.
E-mail address: Xmshen@shsmu.edu.cn (X. Shen).

https://doi.org/10.1016/j.fsigen.2017.11.005
Received 24 August 2017; Received in revised form 15 October 2017; Accepted 11 November 2017
Available online 17 November 2017
1872-4973/

make the estimated value as close to CA as possible. However, the currently available methods are still suboptimal with large uncertainty.

Skeletal and dental ages (SA and DA, respectively) are widely used in age estimation in different populations [7,8]. The combination of SA and DA has recently been recommended because of its higher accuracy than either maturation index alone [9–11]. However, to the best of our knowledge, no confirmation study has been conducted in Chinese children and its applicability to different populations remains uncertain due to potential ethnic variations in skeletal and dental maturation. Moreover, in contrast to CA, biological age (BA), including SA and DA, is influenced by genetic and environmental factors, lifestyle and diseases [12], which decrease the accuracy of the estimate. Newly available technologies may help in this regard.

Previously, various molecular biomarkers have been used to estimate age, including telomere length [13], DNA damage response [14], aspartic acid racemization [15,16], mitochondrial deletions [17], advanced glycation end products [18], T-cell DNA rearrangement [19], and DNA methylation [20]. Unfortunately, none of them, except DNA methylation, have provided an acceptable accuracy that is clinically useful [21]. Several studies reported that DNA methylation markers—age-associated CpG sites—were good predictors for age in adults [22–27]. The methylation statuses of CpG sites were explored using methylome-wide DNA profiling [28]. Bocklandt et al. identified 88 CpG sites in saliva samples of 34 twins and built a regression model with an average accuracy of 5.2 years [29]. Weidner et al. demonstrated a predictive accuracy of 3.34 years with a model based on the methylation levels of 102 age-associated CpG sites in 575 Caucasians [30]. With advancing knowledge of age-associated CpG sites, better accuracy can be achieved in age estimation with fewer markers. Yi et al. identified and validated 16 CpG sites of 8 gene fragments from blood samples in 65 participants. The prediction model explained 91.8% of the variance in age [31]. Xu et al. built a prediction model with an accuracy of 4.7 years based on six CpG sites validated in 50 women [32]. Nevertheless, most relevant studies have focused on the age estimation of adults and explored rates of aging. It is unclear how useful DNA methylation markers are for age estimation in children.

The present study aimed to: (1) identify the most reliable single age estimation measure based on SA (GP, TW3-RUS, and TW3-Carpal) and DA (Demirjian and Willems); (2) identify and validate age-associated CpG sites in children; and (3) improve the accuracy and precision of age estimation by incorporating DNA methylation markers with traditional X-ray techniques (SA and DA estimation methods) in Chinese children.

2. Materials and methods

2.1. Study design

A total of 124 children (78 boys and 46 girls, from 6 to 15 years old) who visited the Child Health Clinic at the Shanghai Children's Medical Center for routine physical examinations and dental check-ups were recruited in the current study using the following inclusion criteria: (1) children who were able to have X-rays taken of left hand-wrist bones (HW-XRs) and orthopantomograms (OPGs), as well as whole blood samples collected on the same day; (2) children who were Han Chinese and had no medical or pathological conditions that might affect growth (e.g., dwarfism and cancer); and (3) children who had no hypodontia or hyperdontia. The gender information and date of birth of the child were reported by the parents. CA was calculated by subtracting the birth date from the measuring date and was expressed in years with two decimals. The mean (standard deviation) of CA was 9.21 (1.73) years for boys, 8.77 (1.66) years for girls, and 9.05 years (1.71 years) for all participants.

2.2. Ethics statement

The Shanghai Children's Medical Center Human Ethics Committee granted ethical approval for this study (SCMCIRB-K2016028), and a written consent was obtained from the parents of each participant.

2.3. X-ray measures

SA was evaluated with three methods: the Greulich and Pyle atlas method (GP) [33], the Tanner and Whitehouse method version 3 on radius-ulna-short bones (TW3-RUS) and on carpal bones (TW3-Carpal) [34]. DA was evaluated by two methods: the Demirjian method [35] and the Willems method [1]. The HW-XRs were rated by two independent radiologists and the OPGs were evaluated by two experienced dentists separately. Approximately 20% of the randomly selected radiographs of HW-XRs and OPGs were re-rated by readers separately chosen from the original specialists after 1 month. The intra- and inter-examiner agreements were measured by intraclass correlation coefficients (ICC). As shown in Table S1, the ICC values of different methods for SA and DA were all more than 0.930, indicating the consistency and reproducibility in age evaluation by the same and different readers.

2.4. DNA sample extraction and preparation

Genomic DNA was extracted from blood samples using the QIAamp DNA Blood Mini Kit (Qiagen, Hilden, Germany) and stored at −20 °C before use, and the purified gDNA was quantified and assessed by NanoDrop ND-2000 (Thermo Scientific, Wilmington, DE, USA). A total of 500 ng DNA was bisulfite converted using EZ DNA Methylation-Gold™ Kits (Zymo Research, Irvine, CA, USA) according to the manufacturer's instructions, following amplification, fragmentation, and precipitation of bisulfite-converted DNA. DNA samples were then re-suspended and allocated onto the Illumina HumanMethylation450 Beadchips.

2.5. Microarray analysis

Profiles were analyzed using the Illumina HumanMethylation450 BeadChip, which can measure the methylation status of 485,577 CpG sites. After resuspension, DNA samples were hybridized using a hybridization oven and the Beadchips were washed out. The BeadChips were imaged and scanned with an iScan reader (Illumina, San Diego, CA, USA). The DNA methylation level at each CpG site on the microarray was detected using GenomeStudio software (Illumina, San Diego, CA, USA) and calculated as a β value, the fraction of methylated cytosines at that location. The β value ranges from 0 to 1, with "0" representing completely unmethylated and "1" representing completely methylated. We have submitted the raw data to the NCBI Gene Expression Omnibus (GEO) under the accession number GSE104812.

2.6. Validation of age-associated CpG sites using droplet digital PCR techniques

The absolute quantification of DNA methylation levels of specific sites from 124 children was performed by droplet digital PCR (ddPCR; QX200, Bio-Rad, Hercules, CA). Primers were designed by the EpiDesigner online tool (http://www.epidesigner.com) for targets using the reverse complementary sequences. The Methylated/Unmethylated Primer/Probe mix was used at final concentrations of 900 and 250 nmol/L, respectively. Each 20-uL PCR reaction was loaded into the 8-channel Droplet Generator Cartridge (Bio-Rad, Hercules, CA). Droplet Generator Oil was then added, and droplets were produced using a QX200 Droplet Generator. The generated droplets were loaded into a 96-well plate and the PCR was run in a C100 Touch™ PCR. The plates were incubated at 95 °C for 10 mins, followed by 40 cycles of 95 °C for 30 s and 60 °C for 60 s and incubation at 98 °C for 10 mins. These plates were read on the Bio-Rad QX200 droplet reader and the analysis was conducted with the QuantaSoft v1.6.6.03 software. All methylation quantification experiments included no template control (NTC) wells,