SWARM™ Targeted Sequencing Measures Epigenetic Age

Wei Guo[1*], Xiaojing Yang[1], Mingda Jin[1], Yap Ching Chew[1], Keith Booher[1], Steve Horvath[2], and Xi Yu Jia[1,3]

[1] Zymo Research Corporation, Irvine, CA, USA; [2] Department of Human Genetics, University of California, Los Angeles, CA, USA; [3] Epimorphy, LLC, Costa Mesa, CA, USA

* Corresponding author's email: wguo@zymoresearch.com

ZYMO RESEARCH

The Beauty of Science is to Make Things Simple

Abstract

Aging is the most important risk factor for many chronic diseases including cancer. DNA methylation is the most reliable molecular marker for aging quantification. However, genome-wide DNA methylation profiling techniques, such as reduced representative bisulfite sequencing and Illumina Bead Arrays, are prohibitively expensive and of poor data quality to be widely used in aging research and intervention. Here we report a robust targeted bisulfite sequencing approach called SWARM™ (Simplified Whole-panel Amplification Reaction Method), which is flexible and low cost, requires relatively low DNA starting material, and increases sample throughput.

Several hundreds of age-associated loci including the published Horvath Clock sites were simultaneously amplified and sequenced. An age-predictive model was built using the elastic net regression of DNA methylation levels of the loci and chronological age of blood DNA samples of over 200 healthy subjects of 18 to 80 years old. Our epigenetic age (DNAge) predictor consists of hundreds of loci and achieved a small median age error. The DNAge of samples that were processed on both sequencing and array platforms highly correlated (r > 0.9). In brief, our versatile SWARM™ technology-based DNAge platform is a very fast and useful tool for human aging quantification, intervention, and monitoring, suitable for various sample types, and applicable in aging studies in mice.

Keywords: DNA methylation, targeted bisulfite sequencing, epigenetic age clock, DNAge, elastic net

The DNAge™ Epigenetic Aging Clock

An epigenetic aging clock is based on measuring natural DNA methylation levels to estimate the biological age of a tissue or cell type. A preeminent example and the golden standard of age prediction is Dr. Steve Horvath's age clock, which is based on 353 human epigenetic markers. The DNAge™ Epigenetic Aging clock builds upon the Horvath's clock and utilizes SWARM™ (Simplified Whole-panel Amplification Reaction Method) technology to analyze DNA methylation patterns of >500 loci, providing epigenetic age predictions in a high throughput manner.

A penalized regression model's coefficients b_0, b_1, …, b_n relate to transformed age as follows

$$F(\text{chronological age}) = b_0 + b_1 CpG_1 + \cdots + b_n CpG_n + \text{error}$$

DNAge™ is estimated as follows

$$DNAge^{TM} = inverse.F(b_0 + b_1 CpG_1 + \cdots + b_n CpG_n)$$

Sequencing Performs Better than Arrays in Epigenetic Age Estimation

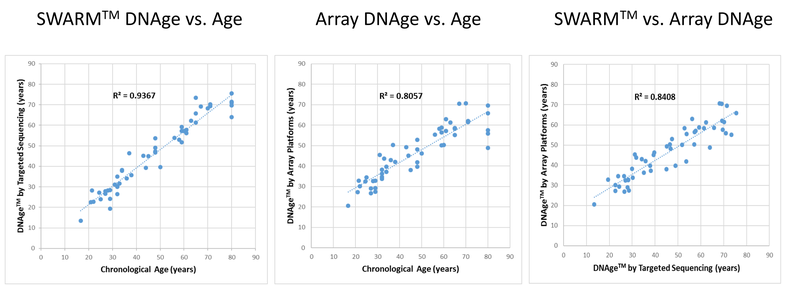

Figure 1. Performance comparison of DNAge estimated by age predictors using methylation values obtained from the same samples (n=51) by the SWARM™ sequencing and the 450K or EPIC methylation arrays. The SWARM™ sequencing-based clock achieved a median test error of 2.0 years while that of the array-based clock was 5.9 years, demonstrating a better performance of the sequencing platform.

DNAge™ Clocks

A DNAge Clock for Whole Blood Samples

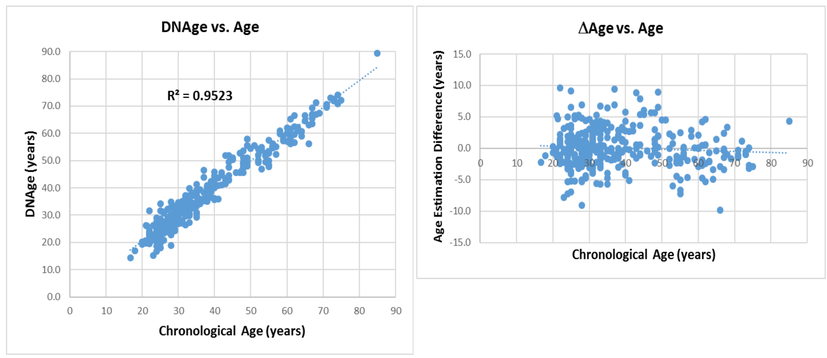

Figure 2. An epigenetic aging clock was built and optimized for whole blood samples. (A) The blood DNAge clock achieved a median absolute age error of 1.9 years meaning that the test error is less than 1.9 years in a half of samples. (B) The test error is consistent in the age range of 20 to 85 years old.

A DNAge Clock for Urine Samples

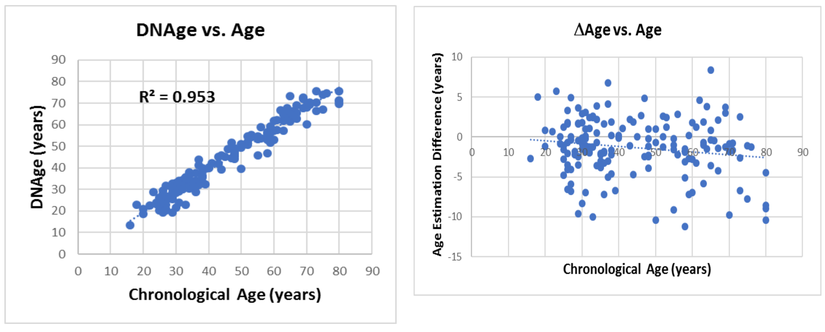

Figure 3. An epigenetic aging clock was built and optimized for non-invasive urine samples. (A) The urine DNAge™ clock achieved a median absolute age error of 2.2 years meaning that the test error is less than 2.2 years in a half of samples. (B) The test error is consistent in the age range of 20 to 80 years old.

Performance Comparison of DNAge™ vs. Horvath Clock

Measurement	Description	Usage	*my* DNAge™ Blood (years)	*my* DNAge™ Urine (years)	Horvath Clock (years)
Median test error	Median absolute difference between DNAge and calendar age. A half of samples had a test error less than this value.	Study whether the DNAge clock is well calibrated.	1.9	2.2	3.6
Age correlation	Pearson correlation coefficient between DNAge and chronological age.	Determine whether the DNAge and calendar age were strongly associated.	r = 0.98	r = 0.96	r = 0.96
Average age acceleration	Average difference between DNAge and chronological age.	Determine whether the DNAge of a given tissue is consistently higher (or lower) than expected.	0.1	-2.0	2.0
Standard deviation	Variation of age acceleration.	Study how disperse age acceleration is.	3.2	4.7	5.3
Variation of repeat tests	Median two standard deviations of difference between DNAge of replicate samples. If a sample is tested multiple times, 95% of DNAge values would be within the mean +- 2SD.	Study how reproducible DNAge test of replicate samples is.	1.7	1.7	3.2

Table 1. The SWARM™ sequencing-based DNAge™ clock performs better than the methylation array-based Horvath clock in epigenetic age estimation. In addition, the sequening platform has higher throughput than the array platform in terms of sample processing.

Nature vs. Nuture

Men Age Faster Than Women in DNAge

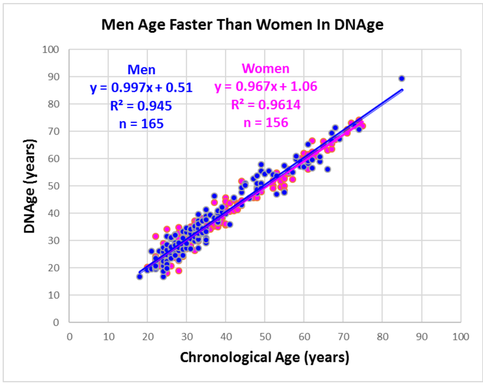

Figure 4. A scatterplot of chronological age vs. DNAge of whole blood samples, showing men age faster than women at a rate of 0.030 years per calendar year. On average, men's DNAge is 0.4 years older than calendar age, while women's DNAge is 0.3 years younger than calendar age. The median age estimation difference of men and women are 0.0 and -0.4 years, respectively, again indicating men age faster than women.

Epigenetic DNAge of Different Human Populations

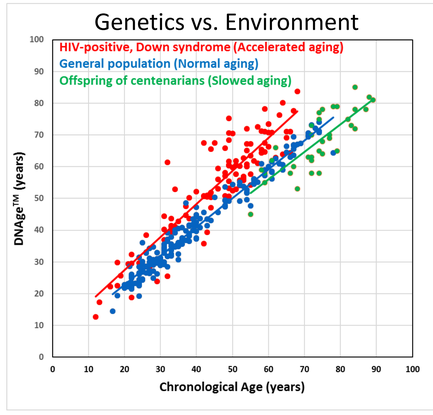

Figure 5. Different human populations show differential aging rates in DNAge, with accelerated aging in blood samples of HIV-positive patients and Down syndrome patients (GSE53840 and GSE52588), normal aging (Zymo Research), and slowed aging (offspring of centenarians, Horvath 2015).

Conclusion

We have developed an improved targeted bisulfite sequencing technology called SWARM™, which is flexible, low cost, requires relatively low DNA starting material, and increases sample throughput. It has been successfully applied in DNA methylation age estimation and achieved better performance and lower cost than methylation array platforms. The DNAge™ epigenetic aging clock is a very fast and useful tool for human aging quantification, intervention and monitoring, is suitable for various sample types, and is applicable in aging studies in mice.