

# 노화 시계를 이용한 알츠하이머병 환자의 후성유전학적 연령 예측

김진영 · 조광원<sup>†</sup>

조선대학교 생명과학과

## Epigenetic Age Prediction of Alzheimer's Disease Patients Using the Aging Clock

Jinyoung Kim, Gwang-Won Cho<sup>†</sup>

*Department of Biology, Chosun University, Gwangju*

### Abstract

Human body ages differently due to environmental, genetic and pathological factors. DNA methylation patterns also differs depending on various factors such as aging and several other diseases. The aging clock model, which uses these differences to predict age, analyzes DNA methylation patterns, recognizes age-specific patterns, predicts age, and grasps the speed and degree of aging. Aging occurs in everyone and causes various problems such as deterioration of physical ability and complications. Alzheimer's disease is a disease associated with aging and the most common brain degenerative disease. This disease causes various cognitive functions disabilities such as dementia and impaired judgment to motor functions, making daily life impossible. It has been reported that the incidence and progression of this disease increase with aging, and that increased phosphorylation of A $\beta$  and tau proteins, which are overexpressed in this disease and accelerates epigenetic aging. It has also been reported that DNA methylation is significantly increased in the hippocampus and entorhinal cortex of Alzheimer's disease patients. Therefore, we calculated the biological age using the Epi clock, a pan-tissue aging clock model, and confirmed that the epigenetic age of patients suffering from Alzheimer's disease is lower than their actual age. Also, it was confirmed to slow down aging.

**Keywords:** Epigenetic clock, Alzheimer's disease, Age acceleration, DNA Methylation

---

(Received May 31, 2023; Revised June 05, 2023; Accepted June 08, 2023)

---

<sup>†</sup> Corresponding author: gwcho@chosun.ac.kr

## 1. 서론

알츠하이머병은 가장 흔한 퇴행성 질환으로 초기에는 해마나 내후각피질 부위에서 국한되어 기억력에서 문제를 일으키다가 두정엽, 전두엽 등을 시작으로 점차 뇌 전체에 전이되어 언어기능이나 판단력 등 여러 인지 기능에 이상이 생기고 운동에도 장애가 생겨 결국 일상 생활이 불가능해지도록 만드는 질병이다. 특히 치매의 원인 중 60-70%가 알츠하이머병이다<sup>[1]</sup>. 전세계적으로 약 5000만명의 사람들이 이 질환을 앓고 있다. 이러한 질병의 원인은 정확히 밝혀져 있지 않지만 유전적 원인, Amyloid 가설, Tau 가설 등 다양한 가설들이 제시되고 있다. 알츠하이머는 질병의 진행은 느리지만 한번 발병하면 다시 되돌릴 수 없다는 것이 가장 큰 문제이다. 그렇기 때문에 치료법이 존재하지 않고 진행을 더 느리게 만드는 완화 효과에 중점을 두고 약이나 식단 조절 등을 사용합니다. 하지만 약의 경우 효과가 미미한 것들이 많고 몇몇 약품들은 심한 부작용을 지니고 있어 약물의 사용 이점이 적습니다. 이전 연구에서는 알츠하이머병이 연령 의존적으로 발병율이 증가하고, 알츠하이머병을 앓는 환자들에게 과발현되는 A $\beta$  또는 Tau 단백질의 인산화가 노화를 촉진시켜 질병의 증상들을 심화시킨다고 보고되었다<sup>[2,3]</sup>.

DNA methylation이 노화의 속도에 영향을 미치는 요인 중 하나로 많이 알려져 있다<sup>[4]</sup>. 이 때문에 DNA methylation과 같은 후성유전학적 변화를 이용하여 노화 마커를 확립하려는 다양한 연구가 진행되어 왔다. 그 중 machine learning을 이용해 DNA methylation pattern을 분석하여 후성유전학적 나이를 예측하는 Epigenetic clock이 확립되었다. DNA methylation은 사람마다 환경적, 유전적, 질병적 등 다양한 요인들에 의해서 다른 pattern을 보인다. Epigenetic clock은 이러한 Pattern을 마커로 이용하여 후성유전학적 연령을 계산한다. 초기에는 하나의 조직만을 사용하여 사용한 조직에서는 매우 높은 정확도로 연령을 예측했지만 사용하지 않은 조직에서는 매우 낮은 정확도를 보였다. 하지

만 이후 multi-tissue나 Pan-tissue를 이용한 추가적인 연구로 인해 사용하지 않은 조직에서도 높은 정확도를 보이는 Epigenetic clock이 제작되었다. 대표적인 예로 Hannum clock, Horvath clock, GrimAge, Epi clock 등이 있다<sup>[5-8]</sup>.

연구 이전에 알츠하이머와 DNA methylation의 연관성을 확인하였다. 질병은 사람의 신체 또는 정신에 문제를 일으키며 나이에 많은 영향을 주고 DNA methylation에도 영향을 준다고 보고되었다<sup>[9]</sup>. 특히 알츠하이머 환자와 정상인의 DNA methylation pattern은 다른 양상을 보인다고 보고되었으며 알츠하이머 환자에게서 DNA methylation이 증가한다고 보고되었다<sup>[10]</sup>. 이러한 DNA methylation의 변화와 정상인과는 다른 DNA methylation의 양상을 통해 epigenetic clock에서도 알츠하이머가 정상인과는 다른 후성유전학적 연령을 보일 것이라고 예측했다. 이를 확인하기 위해 Epi clock을 이용하여 후성유전학적 연령을 계산하였다. 또한 Pathway Analysis를 진행하여 후성유전학적 연령이 어떤 pathway로 이루어졌는지 확인하였다.

## 2. 재료 및 방법

### 2.1. 데이터셋 수집

Gene Expression Omnibus (GEO)를 통해 본 연구에 사용된 모든 데이터 세트들을 수집하였다. 데이터 세트들은 모두 Illumina사의 Infinium Human methylation 450K를 사용하여 추출하였고 IDAT 형식의 파일만을 수집하였다. 또한 노화 가속도를 구하기 위해 실제 연령이 있는 데이터 세트 위주로 수집하였다. 질병과 정상 데이터 모두 존재하는 데이터 세트를 사용하였고 조직의 경우 기억에 관계되어 있으며 알츠하이머병이 초기부터 병변을 보이는 해마와 내후각피질 두 가지를 중심으로 데이터를 수집하였다. 총 사용된 샘플은 230개이고 각 질병의 나이대는 50대 이상의 노년층으로 수집하였다. 표 1은 연구에 사용된 데이터 세트를 정리해 놓은 것이다.

**Table 1.** Cleaning up the collected dataset (Alzheimer's disease & Control).

Dataset ID	Tissue	Technology	No of samples	Data type	Age Range
GSE125895	Hippocampus	Epigenetic array	65	RAW	51.83-99.29
GSE125895	Entorhinal cortex	Epigenetic array	69	RAW	51.83-92.29
GSE105109	Entorhinal cortex	Epigenetic array	96	RAW	58-99

## 2.2 데이터 전처리

R패키지 'minfi'를 이용해 전처리 분위수를 통해 데이터를 정규화하여 Beta value를 계산하였다. 450K array의 총 485512개의 CpG probes를 사용하였다.

## 2.3 후성유전학적 연령 계산

후성유전학적 연령 계산을 위해 Epi clock을 사용하였다. Epi clock은 총 6761개의 CpG site를 사용하며, 범조직을 이용하여 제작된 머신 러닝 모델이다. 제작에 들어간 여러 조직 중 혈액 조직이 가장 많이 사용되어 혈액 조직에서 높은 정확도를 보이고 혈액 샘플 중 전혈 98%로 가장 높은 정확도를 보인다. 또한 해마 조직도 사용되어 95%의 높은 정확도를 보인다. 또한 범조직이 사용되어 제작에 들어가지 않은 몇몇 조직에서도 높은 정확도를 보인다. 사용은 다음과 같다.

우선 Beta value값에서 사용할 6761개의 CpG site를 Python(Anaconda3)를 사용하여 추출하고 Excel을 이용해 Epi clock의 계산식으로 후성유전학적 연령을 계산하였다.

## 2.4 후성유전학적 연령 가속 계산

연령 가속은 Excel을 이용해 후성유전학적 연령에서 실제 나이를 빼 계산하였다. 이는 실제 나이에 비해 얼마나 연령이 가속화 혹은 감속화 되었는지 확인하기 위해 진행하였다.

## 2.5 Differential Methylation Position Analysis

모든 데이터 세트를 질병군과 정상군으로 구분하여 하나의 csv 파일에 구분하여 정리한다. DMPA는

R패키지인 'minfi'를 이용하여 질병과 정상 데이터 세트의 6761개의 CpG site의 methylation 차이를 비교분석하여 유의성을 판별하였다.

## 2.6 Pathway Analysis

DMPA에서 나온 결과를 토대로 상위 10개의 CpG site를 선정 후 CpG site와 연결된 Gene을 추출하였다. 추출한 Gene을 G profiler를 통해 분석하였다.

## 2.7 통계 분석 및 데이터 시각화

정확도는 MAE, RMSE, R square를 엑셀을 통해 계산식을 이용하여 사용하였다. 모든 데이터의 시각화는 GraphPad Prism 8.0.1.244 x64를 사용하여 수행하였다.

## 3. 결과 및 고찰

### 3.1 샘플 특성

샘플은 총 두 개의 데이터 세트를 사용했다. 두 샘플 모두 사후 기증된 뇌를 사용하여 진행되었다. 표 2를 보면 샘플의 조직, 연령, 성별 등 샘플을 추출한 참가자들의 설명적 특성을 나타낸다. 샘플은 알츠하이머병 환자들의 군집인 질병군과 정상인들의 군집인 정상군으로 나뉘고 조직은 해마와 내후각피질로 총 네가지 군으로 나뉜다. 첫번째는 알츠하이머병 환자의 해마로 평균 나이 81.95세로 총 17개의 샘플을 사용했다. 두번째는 정상인의 해마로 평균 나이 61.71세로 총 48개의 샘플을 사용했다. 세번째는 알츠하이머병 환자의 내후각피질로 평균 나이 81.78세로 총 88개의 샘플을 사용했다. 네번째로 정상인의 내후각피질로 평균 나이 67.83세로

**Table 2.** Sample characteristics of Alzheimer’s disease and Control.

Disease	Tissue	Age range	M:F	Mean age	Q1	Q2	Q3	Sample
AD	Hippo	59-99.29	8:9	81.95	78.31	84.11	87.45	17
CTRL	Hippo	51.83-83.64	29:19	61.71	55.61	59.44	64.97	48
AD	ERC	59-97	49:39	81.78	76.7	83	88	88
CTRL	ERC	51.83-99	43:34	67.83	58.11	64.91	77.49	77

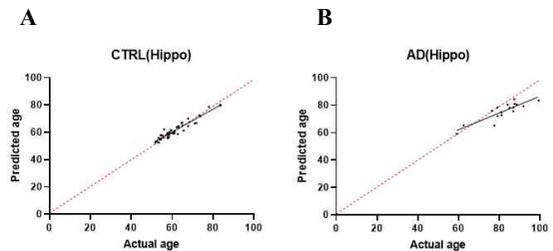
총 77개의 샘플을 사용했다. 각 군은 사분위도의 계산식을 이용해 각 군의 샘플 나이별 분포도를 구하였다. 계산하였는데 전체적으로 정상인 군이 알츠하이머병 환자군에 비해 실제 나이가 낮은 것을 확인할 수 있다. 정상군의 경우 Epi clock에서 얼마나 정확도를 보이는지 확인하기 위해 수행하기 때문에 연령의 차이는 크게 고려되지 않았다.

### 3.2 해마 조직의 후성유전학적 연령

그림 1을 보면 Epi clock에서 정상인의 해마는 clock이 설정한 정상 회귀선과 일치하는 것을 확인하였다. (Fig. 1A) 또한 표 3에서 정상군의 R2값이 0.9116으로 높은 정확도를 보인다. 이는 해마 조직이 Epi clock에서 높은 정확도를 보인다는 것을 암시한다. 반면 알츠하이머의 후성유전학적 연령은 실제 나이에 비해 낮은 연령을 보이고 있고 clock의 회귀선보다 낮은 곳에 위치한다. (Fig. 1B) 또한 표 3에서 알츠하이머병군의 R2값을 보면 0.7616으로 정상군에 비해 낮은 값을 보인다. 이는 알츠하이머병 환자의 생물학적 나이가 정상인에 비해 느리게 흐르고 있다는 것을 의미한다. 표 3을 보면 R2 값을 제외한 Epi clock에서의 다른 정확도를 볼 수 있는데 전체적으로 정상군이 알츠하이머 군에 비해 낮은 값을 보인다는 것을 확인할 수 있다.

**Table 3.** Hippocampal tissue accuracy in Alzheimer’s and Control.

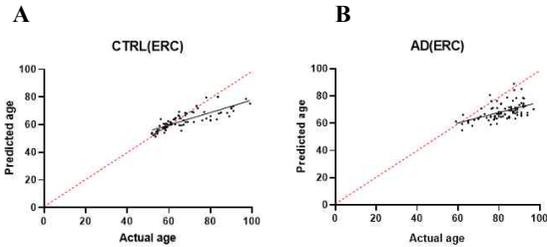
Disease	R2	RMSE	MAE
AD(hippo)	0.7616	8.04	6.71
Ctrl(hippo)	0.9116	2.34	1.87



**Fig. 1.** Comparison of epigenetic age between Alzheimer’s and Control in hippocampal tissue. (A) Control group of Alzheimer’s disease graph. It shows the trend of the epigenetic age of normal people on the Epi clock. Also used to check the accuracy of hippocampal tissue in the Epi clock. (B) Graph of Alzheimer’s disease patients. The epigenetic age of patients with Alzheimer’s disease shows the trend of the Epi clock. (A,B) By comparing the two groups, it is possible to confirm the difference in epigenetic age between normal people and Alzheimer’s patients in the hippocampus.

### 3.3 내후각피질 조직의 후성유전학적 연령

그림 2를 보면 초기 알츠하이머의 병변이 일어나는 다른 조직인 내후각피질의 경우 정상인이 회귀선보다 밑에 있는 것을 볼 수 있는데 이는 Epi clock 내에서 내후각피질의 정확도가 높지 않다는 것을 암시한다. (Fig. 2A) 하지만 알츠하이머병 환자군의 선이 정상인보다 더 낮은 곳에 위치하는 것을 확인할 수 있다. (Fig. 2B) 이는 내후각피질에서도 알츠하이머병 환자의 생물학적 나이가 정상인에 비해 느리게 흐르고 있다는 것을 의미한다. 표 4를 보면 그래프의 정확도를 확인할 수 있다.



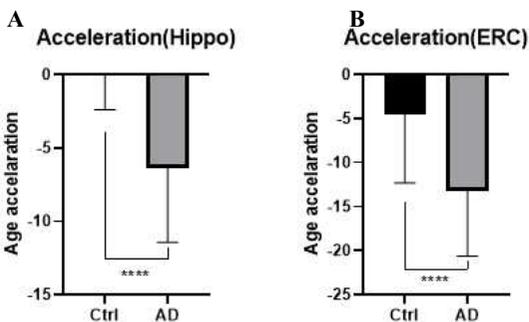
**Fig. 2.** Comparison of epigenetic age between Alzheimer’s and Control in entorhinal cortex tissue. (A) Control of Alzheimer’s disease graph. The epigenetic age trend of normal people is shown in Epi clock. It is also used to check the accuracy of entorhinal cortex tissue in the Epi clock. (B) Alzheimer’s disease patients Graph. The epigenetic age trend of Alzheimer’s disease patients is shown in Epi clock. (A,B) By comparing the two groups, the difference in epigenetic age between normal people and patients with Alzheimer’s disease can be confirmed in the entorhinal cortex.

**Table. 4.** Entorhinal cortex tissue accuracy in Alzheimer’s and Control.

Disease	R2	RMSE	MAE
AD(ERC)	0.2925	15.1	13.5
Ctrl(ERC)	0.6599	8.95	5.8

### 3.4 후성유전학적 연령 가속도

알츠하이머 환자와 정상인의 후성유전학적 가속



**Fig. 3.** Epigenetic age acceleration in the hippocampus and entorhinal cortex. (A) Epigenetic age acceleration in the hippocampus. Showing differences in age acceleration between Alzheimer’s and normal subjects in hippocampal tissue. (B) Accelerated epigenetic age in the entorhinal cortex. Differences in age acceleration between Alzheimer’s and normal subjects in the entorhinal cortex tissue

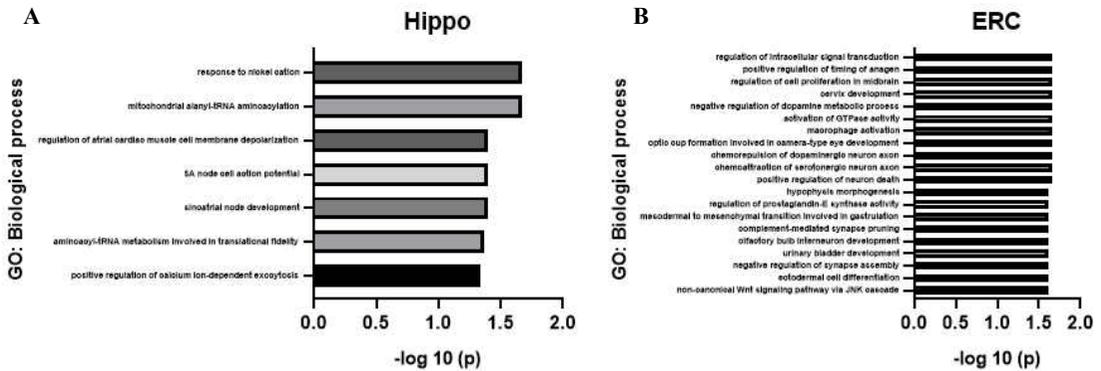
도를 확인하기 위해 후성유전학적 연령 가속도를 확인하였다. 그림 3을 보면 해마와 내후각피질 모두 정상인에 비해 알츠하이머 환자의 노화 가속도가 낮은 것을 확인하였다. (Fig. 3A-B) 이는 Epi clock에서 알츠하이머가 노화를 감속화시킨다는 것을 암시한다.

### 3.5 DMPA

Epi clock에서 각 CpG site의 유의성을 판별하기 위해 DMPA를 진행하였다. 해마 조직과 내후각피질 조직을 각각 DMPA를 진행하여 상위 10개의 CpG site를 추출하였다. 먼저 해마의 DMPA를 통해 나온 상위 10개의 CpG site는 cg18547371, cg11071401, cg11236550, cg26242531, cg13331354, cg12499316, cg07082267, cg07745707, cg07813031, cg04911280 이며 이 중 cg12499316, cg07082267, cg07813031는 유전자를 가지고 있지 않았고 그 외에는 모두 하나 이상의 유전자를 가지고 있었다. 다음으로 DMPA를 통해 나온 내후각피질의 상위 10개의 CpG site는 cg11236550, cg02256631, cg18002862, cg26242531, cg18814344, cg09595163, cg15165154, cg12897164, cg09704168, cg16701059이다. Cg09704168은 유전자를 가지고 있지 않았고 그 외에는 모두 하나의 유전자를 가지고 있었다.

### 3.6 Pathway analysis

그림 4는 각 조직의 상위 10개의 CpG site를 사용하여 나온 Pathway로 중요성이 있는 pathway만을 추출한 것이다. 해마와 내후각피질의 Pathway 개수와 속하는 Pathway가 다른데 이는 같은 질병이라도 조직에 따라 다른 Pathway를 보일 수 있다는 것을 암시한다. 해마의 경우 response to stimulus, metabolic process, biological regulation, multicellular organismal process, developmental process, cellular process 등의 다양한 생물학적 과정이 영향을 주었고 그 중 biological regulation이 가장 많은 비율을 차지했다. (Fig. 4A) 내후각피질의 경우 cellular process, developmental process, metabolic process, biological regulation 등의



**Fig. 4.** Analysis of pathways in the hippocampus and entorhinal cortex. (A) Pathways represented by genes associated with the top 10 CpG sites significantly altered in the hippocampus (B) Pathways represented by genes associated with the top 10 CpG sites significantly altered in the entorhinal cortex

생물학적 과정이 영향을 주었는데 그 중 절반 이상인 11개의 pathway가 developmental process이다. (Fig. 4B) 이를 통해 해마에서는 다양한 생물학적 과정의 영향을 받아 일어난 노화의 감속화를, 그리고 내후각피질에서는 다양한 생물학적 과정 중 Developmental process의 영향을 많이 받아 일어난 노화의 감속화를 Epi clock이 인식하여 후성유전학적 연령이 계산되었다는 것을 암시한다.

#### 4. 결론

결론적으로, 우리는 노화 시계 모델인 Epi clock을 이용하여 후성유전학적 연령 예측을 통해 알츠하이머 환자의 후성유전학적 연령이 실제 연령보다 낮은 것을 확인했다. 또한 알츠하이머 환자의 노화 감속도는 정상인에 비해 더 낮은 것을 확인할 수 있었다. 이는 기존 알려진 알츠하이머가 노화를 촉진해 증상을 심화한다는 관점과는 다른 새로운 관점을 제시한다. 또한 같은 질병이라도 확인하는 조직이 다른 경우 같은 경향을 보인다 하더라도 노화에 영향을 미치는 경로들이 다르다는 것을 확인할 수 있었다. 이는 조직마다 노화의 정도가 다를 수 있다는 것을 암시한다. 하지만 뇌 조직의 샘플들의 경우 대부분 참가자의 사후에

추출되기 때문에 전체적인 나이가 높고 일정하지 않다. 때문에 두 군의 차이를 확인할 수 있으나 그 정확도가 많이 떨어진다는 한계가 존재한다. 또한 해마에서는 높은 정확도를 보이지만 내후각피질의 전체적인 Clock의 정확도가 낮은 것으로 보아 Epi clock에 있는 뇌조직 중 해마를 제외한 높은 정확도를 가진 조직을 선택하여 추가적인 확인이 필요할 것으로 보인다.

#### 5. 참고문헌

- [1] DeTure, M.A. and D.W. Dickson, "The disease. neuropathological diagnosis of Alzheimer's Mol Neurodegener", Vol. 14, No. 1, p. 32, 2019.
- [2] Rajmohan, R. and P.H. Reddy, "Amyloid-Beta and Phosphorylated Tau Accumulations Cause Abnormalities at Synapses of Alzheimer's disease Neurons", J Alzheimers Dis, Vol. 57, No. 4, pp. 975-999, 2017.
- [3] Liu, R.M., Aging, "Cellular Senescence, and Alzheimer's Disease", Int J Mol Sci, Vol. 23, No. 4, 2022.
- [4] Ferrucci, L., et al., "Measuring biological aging in humans: A quest", Aging Cell,

- Vol. 19, No. 2, p.e13080, 2020.
- [5] Hannum, G., et al., *Genome-wide methylation rates. profiles reveal quantitative views of human aging* Mol Cell, Vol. 49, No. 2, p. 359-367, 2013.
- [6] Horvath, S., *DNA methylation age of human tissues and cell types*. Bio Med Center, 2013.
- [7] Ake T. Lu, A.Q., James G. Wilson, Alex P. Reiner, Abraham Aviv, Kenneth Raj, Lifang Hou, Andrea ABaccarelli, Yun Li, James D. Stewart, Eric A.. Whitsel, Themistocles L.Assimes, Luigi Ferrucci, Steve Horvath, *DNA methylation GrimAge strongly predicts lifespan and healthspan*. Aging, 2019.
- [8] Vijayakumar, K.A. and G.W. Cho, *Pan-tissue methylation aging clock: Recalibrated and a method to analyze and interpret the selected features*. Mech Ageing Dev, Vol. 204, p. 111676, 2022.
- [9] Jin, Z. and Y. Liu, *DNA methylation in human diseases*. Genes & Diseases, Vol. 5, No. 1, pp. 1-8, 2018.
- [10] Qazi, T.J., et al., *Epigenetics in Alzheimer's Disease: Perspective of DNA Methylation*. Mol Neurobiol, Vol. 55, No.2, pp. 1026-1044, 2018.