Journal of Clinical Lipidology (2025) 000, 1–2

Journal of Clinical Lipidology

Letter to the Editor

Enhancing lipoprotein(a) association studies: A complementary approach to principal component analysis

Kempkes et al. have made a valuable contribution to the field of cardiovascular medicine by investigating the link between lipoprotein(a) (Lp[a]), atherosclerotic cardiovascular disease (ASCVD), and human monocyte inflammation.[1] In their observational study of 60 subjects, they demonstrated that individuals with high Lp(a) levels (>350 nmol/L) exhibit an inflammatory signature, including significantly higher plasma concentrations of the cytokine C-X-C motif chemokine ligand 10 (CXCL10), compared to those with low Lp(a) levels (<25 nmol/L). Nevertheless, their conclusion that this inflammation is not accompanied by changes in chromatin accessibility in circulating classical monocytes warrants further discussion. The methodological approach, particularly the use of principal component analysis (PCA) and bulk sequencing, may be limited in its ability to capture the complex and nonlinear characteristics of epigenetic data.

This interpretation relies heavily on PCA, a linear dimensionality reduction technique. Although PCA is widely used for visualizing and interpreting high-dimensional datasets, its effectiveness depends on the assumption that the underlying data structure is linear.[2-6] However, this assumption may fall short in capturing the complexity of biological systems, where epigenetic regulation is shaped by gene-environment interactions and dynamic cellular responses that often follow nonlinear patterns. As a result, linear methods may fail to detect subtle but clinically relevant signals, potentially leading to oversimplified interpretations of the epigenetic basis of inflammation.

While Kempkes et al. included additional analyses, their approach to dimensionality reduction and visualization did not fully account for the nonlinear structure of epigenetic dynamics. Combined with bulk sequencing, this may have limited their ability to detect subtle temporal shifts and cellular interdependencies that potentially precede inflammation. Although nonlinear visualization tools like t-Distributed Stochastic Neighbor Embedding (t-SNE) and Uniform Manifold Approximation and Projection (UMAP) are widely used in high-dimensional biological data analysis, they remain primarily exploratory and are not well-suited for feature prioritization or statistical inference.

To address these limitations, future studies should adopt a more comprehensive analytical framework that focuses on biologically meaningful feature selection and robust statistical modeling. Unsupervised methods such as Feature Agglomeration and Highly Variable Gene Selection offer promising alternatives to traditional linear approaches, enabling the identification of key microbial signatures without imposing restrictive assumptions.[7,8] Additionally, nonparametric correlation measures like Spearman's rho and Kendall's tau are well-suited for detecting monotonic relationships in epigenetic data, thereby enhancing both analytical precision and interpretability.[9,10]

In conclusion, advancing epigenetic research in cardiovascular medicine requires analytical frameworks that move beyond linear assumptions. Robust approaches such as Feature Agglomeration and Highly Variable Gene Selection allow for biologically meaningful feature selection, while nonparametric correlation measures like Spearman's rho and Kendall's tau offer greater flexibility in capturing complex associations. These methods provide a more accurate basis for identifying early epigenetic indicators of inflammation and improving clinical decision-making.

CRediT authorship contribution statement

Souichi Oka: Conceptualization, Writing – original draft. **Nobuko Inoue:** Investigation. **Yoshiyasu Takefuji:** Project administration, Supervision, Writing – review & editing.

Funding source

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

https://doi.org/10.1016/j.jacl.2025.08.021

Data availability

No new data were generated or analyzed in support of this research.

Souichi Oka*, Nobuko Inoue
Science Park Corporation, Kanagawa, Japan (Drs Oka and Inoue)

Yoshiyasu Takefuji
Faculty of Data Science, Musashino University, Tokyo, Japan (Dr Takefuji)

https://doi.org/10.1016/j.jacl.2025.08.021

*Corresponding author at: Science Park Corporation, 3-24-9 Iriya-Nishi Zama-shi, Kanagawa 252-0029, Japan.
E-mail address: souichi.oka@sciencepark.co.jp (S. Oka)

Received August 22, 2025
Accepted August 25, 2025

References

1. Kempkes RWM, Kraaijenhof JM, van Os BW, et al. Human monocyte inflammation in patients with Lp(a) or atherosclerotic cardiovascular disease is not accompanied by changes in chromatin accessibility in circulating classical monocytes. *J Clin Lipidol*. 2025. doi:10.1016/j.jacl.2025.08.003.
2. Mohseni N, Elhaik E. Biases of principal component analysis (PCA) in physical anthropology studies require a reevaluation of evolutionary insights. *eLife*. 2024;13:RP94685. doi:10.7554/eLife.94685.2.
3. Yao Y, Ochoa A. Limitations of principal components in quantitative genetic association models for human studies. *eLife*. 2023;12:e79238. doi:10.7554/eLife.79238.
4. Elhaik E. Principal component analyses (PCA)-based findings in population genetic studies are highly biased and must be reevaluated. *Sci Rep*. 2022;12(1):14683. doi:10.1038/s41598-022-14395-4.
5. Lenz M, Müller FJ, Zenke M, Schuppert A. Principal components analysis and the reported low intrinsic dimensionality of gene expression microarray data. *Sci Rep*. 2016;6:25696. doi:10.1038/srep25696.
6. Dey R, Lee S. Asymptotic properties of principal component analysis and shrinkage-bias adjustment under the generalized spiked population model. *J Multivar Anal*. 2019;173:145–164. doi:10.1016/j.jmva.2019.02.007.
7. Zhang J, Wu X, Hoi SCH, Zhu J. Feature agglomeration networks for single stage face detection. *Neurocomputing*. 2020;380:180–189. doi:10.1016/j.neucom.2019.10.087.
8. Xie Y, Jing Z, Pan H, et al. Redefining the high variable genes by optimized LOESS regression with positive ratio. *BMC Bioinformatics*. 2025;26:104. doi:10.1186/s12859-025-06112-5.
9. Okoye K, Hosseini S. Correlation tests in R: Pearson Cor, Kendall's tau, and Spearman's rho. In: Okoye K, Hosseini S, eds. *R programming: statistical data analysis in research* Springer Nature; 2024:247–277. doi:10.1007/978-981-97-3385-9_12.
10. Yu H, Hutson AD. A robust Spearman correlation coefficient permutation test. *Commun Stat Theory Methods*. 2024;53:2141–2153. doi:10.1080/03610926.2022.2121144.