Plasma proteins as potential biomarker of aging of single tissue and cell type

Daigo Okada[1*]

[1*]Center for Genomic Medicine, Graduate School of Medicine, Kyoto University, 53 Syogoin-Kawaramachi, Sakyo-ku, Kyoto, 606-8507, Kyoto, Japan.

Corresponding author(s). E-mail(s): dokada@genome.med.kyoto-u.ac.jp;

Abstract

Plasma proteins serve as biomarkers of aging and various age-related diseases. While a number of plasma proteins have been identified that increase or decrease with age, the interpretation of each protein is challenging. This is due to the nature of plasma, which is a mixture of factors secreted by many different tissues and cells. Therefore, the catalog of age-related proteins secreted by a single cell type in a single tissue would be useful for understanding tissue-specific aging patterns. In this study, the author addressed this challenge by integrative data mining of the Human Protein Atlas (HPA) and the recently published result of large-scale aging proteomics research. Finally, we identified the 17 age-related proteins produced by a single tissue and a single cell type: MBL2 and HP in the liver (hepatocytes), SFTPC in the lung (type II alveolar cells), PRL and POMC in the pituitary (anterior cells), GCG, CUZD1 and CPA2 in the pancreas (pancreatic cells), MYBPC1 in skeletal muscle (myocytes), PTH in the parathyroid gland (glandular cells), LPO and AMY1A in the salivary gland (glandular cells), INSL3 in the male testis (Leydig cells), KLK3 and KLK4 in the male prostate (glandular cells), MPO and ACP5 in immune cells. This list of proteins would be potentially useful for understanding age-related changes in the plasma proteome and inter-tissue networks.

Keywords: computational biology, proteomics, aging, biomarker

1 Introduction

Plasma proteins serve as biomarkers of aging and various age-related diseases [1–3]. Plasma contains various protein factors that are secreted by a variety of tissues or cells, and these factors are the primary means of communication between tissues throughout the body. During aging, the level of many plasma factors changes dynamically [3], and these changes reflect the state of tissues throughout the body. The plasma proteome can even predict chronological age with high accuracy [3, 4]. Thus, the plasma proteome provides important information about the aging status of multiple tissues throughout the body.

While a number of plasma proteins have been identified that increase or decrease with age [3], the interpretation of each protein is challenging. This is due to the nature of plasma, which is a mixture of factors secreted by many different tissues and cells. Each tissue is composed of numerous cell types, each of which secretes specific proteins into the plasma. This further complicates the interpretation and tissue annotation of plasma protein markers.

In addition, recent aging research has focused on the dysfunctional communication between tissues that occurs with aging [5]. For a deep understanding of the inter-tissue relationships in aging, it is necessary to understand the aspects of aging in each tissue individually. Among the many proteins known to change with age, identifying the proteins secreted by a single tissue and single cell type should provide useful information for interpreting age-related changes in the plasma proteome and for a systematic understanding of tissue aging.

In this study, the author addressed this challenge by integrative data mining of the Human Protein Atlas (HPA) and the recently published result of large-scale aging proteomics research. As a result, we identified 17 plasma aging biomarkers for a single tissue or single cell type. Because these are believed to be derived from a single tissue or cell, they may reflect the aging status of that particular tissue or cell. It is hoped that this study will provide candidate protein markers that reflect the aging state of a specific tissue or cell population and facilitate understanding of the individual aging system as a whole.

2 Method

The purpose of this study is to identify plasma aging biomarkers that are specific to a particular tissue/cell type (Fig.1). Tissues contain many cell types. Some of these cell types are common to multiple tissues, while others are unique to a particular tissue. Therefore, if multiple tissues and cell types secrete the same proteins into the plasma, the biological interpretation of their aging changes becomes difficult. Therefore, the catalog of age-related proteins secreted by a single cell type in a single tissue would be useful for understanding tissue-specific aging patterns.

The author integrated the Human Protein Atlas and recently published aging proteomics results to identify tissue- and cell-type-specific plasma aging biomarkers.

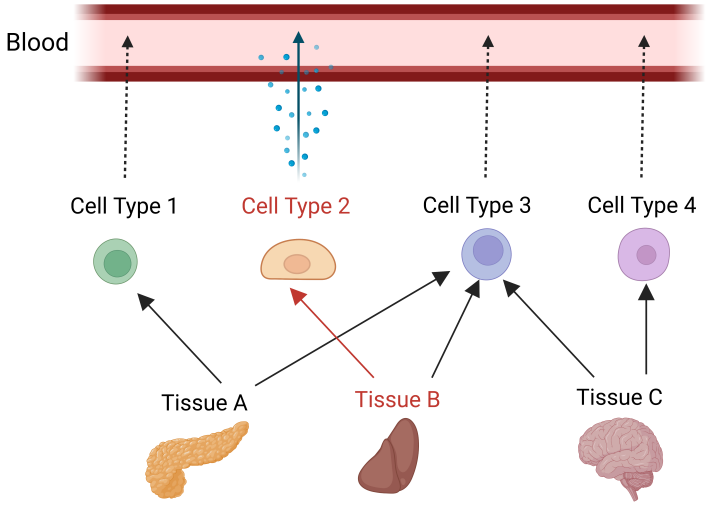

Fig. 1 The concept of the potential plasma biomarker of aging of single tissue and cell type. In this figure, only Tissue B contains Cell type 2. If only Cell type 2 produces a protein and it is observed in plasma, this aging change in plasma levels is a potential biomarker reflecting the aging status of Cell type2 in Tissue B.

The results of the large-scale aging proteomic analysis by Lehallier et al. were used as a list of age-related proteins [3]. This study has examined 2925 protein expressions in over 4000 human subjects and identified the 1,379 proteins with a significant linear relationship with age (FDR Q value < 0.05). This protein list was used for the age-related protein list in our study.

To identify the tissue and cell type specific expressed proteins, the author used the protein expression dataset in human tissues using immunohistochemistry of tissue microarrays from Human Protein Atlas [6]. The data set "normal_tissue.tsv.zip" was obtained from the download page (https://www.proteinatlas.org/about/download). This data contains protein expression information for 145 cell types in 64 human tissues. Data with "Enhanced" reliability scores were extracted and used in subsequent analyses.

Based on this public dataset, a protein expression matrix was constructed for each tissue and cell type. First, since the expression level is expressed in categories, it was converted to a numerical value. Based on the rules proposed in the previous study [7], the categories "High", "Medium", and "Low" were converted to 3, 2, and 1, and "Ascending", "Descending", and "Not representative" to 1, respectively. "Not detected" was converted to 0. If there were multiple data for the same ensemble gene ID and tissue (or cell type), the maximum value was taken as the expression level of that protein in that tissue (or cell type). The elements of the matrices for which there were no corresponding measurements were also assigned -1 as a missing value. Finally, we

constructed the two consensus expression matrices of proteins × tissues and proteins × cell types.

Cell type specific expressed protein and tissue specific expressed protein were identified by the following procedures. Proteins with an expression level of 3 in only one tissue and ≤ 0 in all other tissues were defined as proteins specifically expressed in that tissue. Proteins that did not meet the above conditions were annotated as "not specific". Following the same rule, labels for cell type specificity were also annotated for each protein. As a result, each protein had two annotations of tissue and cell type specificity.

Next, the re-annotation of age-related plasma proteins using this list of proteins expressed specifically in a tissue or cell type was performed as follows. The Entrez Gene IDs of each age-related protein were converted to Ensemble Gene ID using the ID table obtained from Ensemble Biomart [8] for integration with information from the Human Protein Atlas. Tissue and cell type specificity labels were annotated for each age-related protein. In the case that an age-related protein corresponds to multiple ensemble gene IDs and multiple annotations, "not specific" was given priority.

Finally, age-related proteins with both cell type-specific and tissue-specific annotations were extracted. A literature search was performed on these proteins to identify age-related protein markers of a single tissue and cell type.

3 Result and Discussion

A total of 17 proteins were detected as potential age-related proteins with both cell type and tissue specificity (Table 1). For example, MBL2 and HP are detected in hepatocytes of the liver. Hepatocytes in the liver are considered a major source of serum protein [9]. MBL2 is negatively associated with age and the polymorphisms associated with low serum MBL2 have been reported to be associated with several liver diseases [10–12]. HP correlates positively with age and is known to be an inflammatory marker [13]. These proteins may be biomarkers for measuring the biological age of hepatocytes in the liver. Similarly, the following proteins may also serve as biomarkers of tissue aging. SFTPC in the lung (type II alveolar cells), PRL and POMC in the pituitary (anterior cells), GCG, CUZD1, and CPA2 in the pancreas (pancreatic cells), MYBPC1 in skeletal muscle (myocytes), PTH in parathyroid (gland cells), LPO and AMY1A in salivary gland (gland cells). These proteins may be useful as biomarkers of aging in specific tissues.

On the other hand, it is difficult to determine whether some proteins associated with male tissues are age-related plasma proteins produced in a single tissue and cell type. One age-related plasma protein has been identified in the testis: INSL3 in testis Leydig cells. This is already known as a Leydig cell biomarker and can be used as a biomarker for various age-related diseases [14]. Actually, in females, INSL3 is produced by the theca interna cells of the antral follicles [15], and is not strictly a testis-specific protein. This is because the theca interna cell is not included in the dataset. As another male tissue, KLK3

and KLK4 were identified in the prostate (glandular cells). These proteins would not necessarily be prostate-specific as they have also been detected in female blood [16, 17]. One of the limitations of this study is that some cell types and tissues are not fully included in the HPA dataset, resulting in the detection of proteins produced in multiple tissues. Nevertheless, at least in males, these proteins are a potential biomarker of aging of the testis or prostate.

Similarly, proteins specific to immune cells should be interpreted with caution. For example, in bone marrow, MPO has been identified in hematopoietic cells. MPO is a gene expressed in neutrophils and its level in plasma is considered a potential neutrophil marker [18, 19]. Therefore, this protein may be considered a specific plasma protein of neutrophils rather than hematopoietic cells in the bone marrow. In any case, potentially useful as an aging biomarker of immune function. Similarly, ACP5 identified in the lung (macrophage) is also controversial. Immune cells such as macrophages are present in many tissues throughout the body. These proteins may be useful as markers of aging for specific immune cell subsets rather than specific tissues.

4 Conclusion

In this study, public data were re-analyzed to comprehensively search for plasma proteins of aging biomarkers of single tissue and cell type. Finally, we identified the age-related proteins produced by a single tissue and a single cell type: MBL2 and HP in the liver (hepatocytes), SFTPC in the lung (type II alveolar cells), PRL and POMC in the pituitary (anterior cells), GCG, CUZD1 and CPA2 in the pancreas (pancreatic cells), MYBPC1 in skeletal muscle (myocytes), PTH in the parathyroid gland (glandular cells), LPO and AMY1A in the salivary gland (glandular cells), INSL3 in the male testis (Leydig cells), KLK3 and KLK4 in the male prostate (glandular cells), MPO and ACP5 in immune cells. This list of proteins would be potentially useful for understanding age-related changes in the plasma proteome and inter-tissue networks.

Acknowledgments. Fig.1 was created using BioRender (https://www.biorender.com/).

Declarations

Funding

Not applicable.

Conflict of interest

The author has no conflict of interest to be declared.

Ethics approval

Not applicable.

Data and Code availability

This study used public data only. The R code used in this study is available at (https://github.com/DaigoOkada/aginghpa).

References

[1] Lu, J., Huang, Y., Wang, Y., Li, Y., Zhang, Y., Wu, J., Zhao, F., Meng, S., Yu, X., Ma, Q., *et al.*: Profiling plasma peptides for the identification of potential ageing biomarkers in chinese han adults. PLoS One **7**(7), 39726 (2012)

[2] Xu, R., Gong, C., Duan, C., Huang, J., Yang, G., Yuan, J., Zhang, Q., Xiong, X., Yang, Q.: Age-dependent changes in the plasma proteome of healthy adults. The journal of nutrition, health & aging **24**, 846–856 (2020)

[3] Lehallier, B., Gate, D., Schaum, N., Nanasi, T., Lee, S.E., Yousef, H., Moran Losada, P., Berdnik, D., Keller, A., Verghese, J., *et al.*: Undulating changes in human plasma proteome profiles across the lifespan. Nature medicine **25**(12), 1843–1850 (2019)

[4] Enroth, S., Enroth, S.B., Johansson, Å., Gyllensten, U.: Protein profiling reveals consequences of lifestyle choices on predicted biological aging. Scientific reports **5**(1), 1–10 (2015)

[5] Tian, Y.E., Cropley, V., Maier, A.B., Lautenschlager, N.T., Breakspear, M., Zalesky, A.: Heterogeneous aging across multiple organ systems and prediction of chronic disease and mortality. Nature Medicine, 1–11 (2023)

[6] Uhlén, M., Fagerberg, L., Hallström, B.M., Lindskog, C., Oksvold, P., Mardinoglu, A., Sivertsson, Å., Kampf, C., Sjöstedt, E., Asplund, A., *et al.*: Tissue-based map of the human proteome. Science **347**(6220), 1260419 (2015)

[7] Prakash, A., García-Seisdedos, D., Wang, S., Kundu, D.J., Collins, A., George, N., Moreno, P., Papatheodorou, I., Jones, A.R., Vizcaíno, J.A.: Integrated view of baseline protein expression in human tissues. Journal of Proteome Research (2022)

[8] Martin, F.J., Amode, M.R., Aneja, A., Austine-Orimoloye, O., Azov, A.G., Barnes, I., Becker, A., Bennett, R., Berry, A., Bhai, J., *et al.*: Ensembl 2023. Nucleic Acids Research **51**(D1), 933–941 (2023)

[9] Tada, T., Ohkubo, I., Niwa, M., Sasaki, M., Tateyama, H., Eimoto, T.: Immunohistochemical localization of zn-alpha 2-glycoprotein in normal human tissues. Journal of Histochemistry & Cytochemistry **39**(9), 1221–1226 (1991)

[10] Vasconcelos, L.R.S., Fonseca, J.P.L., do Carmo, R.F., de Mendonça, T.F., Pereira, V.R.A., Lucena-Silva, N., Pereira, L.M.M.B., Moura, P., de Mendonça Cavalcanti, M.d.S.: Mannose-binding lectin serum levels in patients with leprosy are influenced by age and mbl2 genotypes. International Journal of Infectious Diseases **15**(8), 551–557 (2011)

[11] Alves Pedroso, M., Boldt, A., Pereira-Ferrari, L., Steffensen, R., Strauss, E., Jensenius, J.C., Ioshii, S., Messias-Reason, I.: Mannan-binding lectin mbl2 gene polymorphism in chronic hepatitis c: association with the severity of liver fibrosis and response to interferon therapy. Clinical & Experimental Immunology **152**(2), 258–264 (2008)

[12] Koutsounaki, E., Goulielmos, G.N., Koulentaki, M., Choulaki, C., Kouroumalis, E., Galanakis, E.: Mannose-binding lectin mbl2 gene polymorphisms and outcome of hepatitis c virus-infected patients. Journal of clinical immunology **28**, 495–500 (2008)

[13] Maffei, M., Funicello, M., Vottari, T., Gamucci, O., Costa, M., Lisi, S., Viegi, A., Ciampi, O., Bardi, G., Vitti, P., *et al.*: The obesity and inflammatory marker haptoglobin attracts monocytes via interaction with chemokine (cc motif) receptor 2 (ccr2). BMC biology **7**, 1–14 (2009)

[14] Ivell, R., Heng, K., Severn, K., Antonio, L., Bartfai, G., Casanueva, F.F., Huhtaniemi, I.T., Giwercman, A., Maggi, M., O'Connor, D.B., *et al.*: The leydig cell biomarker insl3 as a predictor of age-related morbidity: Findings from the emas cohort. Frontiers in Endocrinology **13**, 2848 (2022)

[15] Ivell, R., Anand-Ivell, R.: Insulin-like peptide 3 (insl3) is a major regulator of female reproductive physiology. Human reproduction update **24**(6), 639–651 (2018)

[16] Planque, C., Li, L., Zheng, Y., Soosaipillai, A., Reckamp, K., Chia, D., Diamandis, E.P., Goodglick, L.: A multiparametric serum kallikrein panel for diagnosis of non–small cell lung carcinoma. Clinical cancer research **14**(5), 1355–1362 (2008)

[17] Raimondo, S., Gentile, M., Esposito, G., Gentile, T., Ferrara, I., Crescenzo, C., Palmieri, M., Cuomo, F., De Filippo, S., Lettieri, G., *et al.*: Could kallikrein-related serine peptidase 3 be an early biomarker of environmental exposure in young women? International Journal of Environmental Research and Public Health **18**(16), 8833 (2021)

[18] Ramachandra, C.J., Ja, K.M.M., Chua, J., Cong, S., Shim, W., Hausenloy, D.J.: Myeloperoxidase as a multifaceted target for cardiovascular protection. Antioxidants & redox signaling **32**(15), 1135–1149 (2020)

[19] Grisham, M.B., Everse, J., Janssen, H.F.: Endotoxemia and neutrophil activation in vivo. American Journal of Physiology-Heart and Circulatory Physiology **254**(5), 1017–1022 (1988)