論　文　内　容　要　旨

RNA-Sequencing Analysis Identification of Potential Biomarkers for Diagnosis of Sarcopenia

（RNA 配列解析によるサルコペニア診断のための潜在的バイオマーカーの同定）

Journals of Gerontology Series A biological sciences and medical sciences, 2023, in press.

主指導教員：中野　由紀子 教授
（医系科学研究科　循環器内科学）
副指導教員：北川 知郎 講師
（医系科学研究科　循環器内科学）
副指導教員：石田 万里　准教授
（医系科学研究科　心臓血管生理学）

古谷　　元樹

（医系科学研究科　医歯薬学専攻）

Background Sarcopenia is a geriatric disease that is defined as a combination of low skeletal muscle mass and low muscle function and is associated with increased mortality and disability. Early diagnosis and intervention are required to prevent it. Messenger RNA (mRNA) sequencing (RNA-seq) using blood samples will likely be useful for developing easily accessible and non-invasive biomarkers.

Methods Sarcopenia was diagnosed according to the Asian Working Group for Sarcopenia 2019 consensus. Whole-blood RNA-seq data and associated clinical data were obtained from the National Center for Geriatrics and Gerontology Biobank, Japan. The clinical data were assessed with logistic regression, adjusting for age, sex, and body mass index (BMI). The RNA-seq data were used to detect differentially expressed genes (DEGs) between sarcopenia and normal elderly controls (NCs). To gain further insight into the biological functions of the DEGs, we performed gene ontology (GO)-term enrichment analysis and protein-protein interaction (PPI) network analysis. We also performed immune cell analysis to investigate the differences in immune cell–type composition between phenotypes by using RNA-seq data. A best risk-prediction model was constructed on the basis of a random forest classifier from any combination of the detected biomarker candidates and clinical information (age, sex, and BMI).

Results A total of 114 elderly, consisting of 52 sarcopenia patients (42.30% female, mean age 79.90 ± 5.95 years) and 62 NCs (87.10% female, mean age 75.15 ± 6.22 years), were enrolled. The clinical data analysis showed a significant decrease in stride length in sarcopenia patients (Bonferroni corrected $P = 0.012$). The RNA-seq transcriptome analyses detected six DEGs (*FAR1*, *GNL2*, *HERC5*, *MRPL47*, *NUBP2*, and *S100A11*) as biomarker candidates for sarcopenia diagnosis. Subsequent GO-term enrichment analysis and PPI network analysis using RNA-seq data further detected two functional modules (i.e., hub genes, *MYH9* and *FLNA*) as biomarker candidates. By using any combination of the nine candidates detected and clinical information (age and sex), risk-prediction models were constructed. The best model was constructed by using a combination of clinical information and four potential biomarkers (stride length, *HERC5*, *S100A11*, and *FLNA*), which achieved a high area under the curve (AUC) of 0.91 in a validation cohort (95% confidence interval [CI]: 0.79 to 0.96). When BMI was added, the model achieved a high AUC of 0.95 (95% CI: 0.84 to 0.98). Our risk-prediction models gave significantly higher AUCs than those constructed with only clinical information (Welch's *t*-test $P < 0.001$).

Discussions

Age-related changes in muscle tissue morphology and function, including myocyte apoptosis, mitochondrial dysfunction, and oxidative stress, contribute to the pathogenesis of sarcopenia. Considering this multidimensionality of the disease, the identification of useful

biomarkers is challenging. To address this issue, we used a comprehensive approach, integrating clinical data analysis, RNA-seq transcriptome analysis, GO-term enrichment analysis, PPI analysis, and immune cell analysis. We finally identified four potential biomarkers (stride length, *HERC5*, *S100A11*, and *FLNA*) and constructed risk-prediction models by using a combination of clinical information (age, sex, and BMI) and the biomarkers, thus achieving high AUCs.

BMI is useful for evaluating nutritional status and assessing obesity, which leads to a state of low-grade systemic inflammation. Recently, sarcopenic obesity, defined as a phenotype of the coexistence of sarcopenia and obesity, has received much attention. It carries higher risks of mortality and disability than sarcopenia or obesity alone. BMI remains relatively unchanged in elderly persons because muscle mass gradually declines with aging and there is a relative increase in fat content. It is still controversial as to whether BMI is useful for assessing sarcopenia or obesity in the elderly. We therefore constructed risk-prediction models both with and without BMI. Interestingly, our final risk-prediction models were constructed by using the same four biomarkers. These results imply that the biomarkers that we found have the potential to be robust in sarcopenia diagnosis.

One biomarker detected in our clinical data analysis—stride length—was highly correlated with gait speed, which is used in the diagnosis of sarcopenia (Pearson's $r = 0.91$). We expect that the use of this biomarker will help improve the accuracy of clinical diagnosis. A previous meta-analysis reported that stride length can predict adverse clinical events such as physical disability and mortality. The three remaining biomarkers, which were genes detected in our systems biology analyses by using DEGs obtained from the RNA-seq transcriptome analysis, are associated with inflammation. *HERC5* (*HECT and RLD domain containing E3 ubiquitin protein ligase 5*) acts as a modulator of the antiviral immune response, and its expression is regulated by interferon, inflammatory cytokine interleukin 1 beta, and tumor necrosis factor-alpha. This gene is upregulated in dermatomyositis, a long-term inflammatory disorder that affects the skin and muscles. *S100A11* (*S100 calcium-binding protein A11*) plays essential roles in different biological functions, including enzyme activity, cell growth regulation, tumor progression, and the development of inflammatory diseases such as osteoarthritis and rheumatoid arthritis. The other investigation reported the possibility that *S100A11* is a biomarker of inflammatory myopathies. *FLNA* (*filamin A*), which encodes an actin-binding protein, is necessary for T-cell activation, which is associated with diseases such as the chronic inflammatory disease atherosclerosis. Previous studies have reported that inflammation is an important contributor to the pathogenesis of sarcopenia. We verified that these genes are expressed in

myogenic cells as well as in blood. Our results suggest that the biomarkers identified here not only can be useful for the diagnosis of sarcopenia but also could be directly associated with the pathogenesis of the disease.

Conclusions We have discovered potential biomarkers for early diagnosis of sarcopenia. Further refinement may lead to their future practical use in the clinic.