Matrix metalloproteinase-degraded type I collagen is associated with APOE/TOMM40 variants and preclinical dementia

Man-Hung Eric Tang, PhD, Joseph P.M. Blair, MSc, Cecilie Liv Bager, PhD, Anne-Christine Bay-Jensen, PhD, Kim Henriksen, PhD, Claus Christiansen, MD, and Morten Asser Karsdal, PhD

Neurol Genet 2020;6:e508. doi:10.1212/NXG.0000000000000508

Abstract

Objective
Dysregulation of type I collagen metabolism has a great impact on human health. We have previously seen that matrix metalloproteinase–degraded type I collagen (C1M) is associated with early death and age-related pathologies. To dissect the biological impact of type I collagen dysregulation, we have performed a genome-wide screening of the genetic factors related to type I collagen turnover.

Methods
Patient registry data and genotypes have been collected for a total of 4,981 Danish postmenopausal women. Genome-wide association with serum levels of C1M was assessed and phenotype-genotype association analysis performed.

Results
Twenty-two genome-wide significant variants associated with C1M were identified in the APOE-C1/TOMM40 gene cluster. The APOE-C1/TOMM40 gene cluster is associated with hyperlipidemia and cognitive disorders, and we further found that C1M levels correlated with tau degradation markers and were decreased in women with preclinical cognitive impairment.

Conclusions
Our study provides elements for better understanding the role of the collagen metabolism in the onset of cognitive impairment.

From the ProScion (M.-H.E.T., J.P.M.B., C.L.B.), Herlev; Faculty of Health and Medical Sciences (J.P.M.B.), University of Copenhagen; ImmunoScience (A.-C.B.-J., C.C., M.A.K.), and Endocrinology (K.H.), Nordic Bioscience, Biomarkers and Research, Herlev, Denmark.

Go to Neurology.org/NG for full disclosures. Funding information is provided at the end of the article.

The Article Processing Charge was funded by the authors.

This is an open access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND), which permits downloading and sharing the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.
Many pathologies emerge after menopause, affecting the quality and duration of women’s lives. Metabolic and endocrine changes occurring during the menopause transition have been linked with an increased incidence of chronic inflammatory and autoimmune disorders, in particular increased neuroinflammation. A dysregulated extracellular matrix (ECM) metabolism is a common denominator in several age-related fibroproliferative pathologies and attributes to almost 45% of all deaths in the developed world. Thus, improved preventive and predictive strategies for chronic fibroproliferative diseases could improve both quality of life and enhance longevity.

Type I collagens are among the most abundant extracellular matrix proteins in the body and are expressed in most connective tissues. During homeostasis, type I collagen is maintained in a delicate equilibrium between protein formation and degradation. Remodeling and repair of tissues is therefore essential for sustaining a healthy tissue or organ. During the development of inflammatory diseases, the equilibrium between type I collagen formation and degradation is shifted, leading to an altered tissue remodeling and repair that in turn drives the disease.

Blood-based biomarkers of collagen metabolism have been used as an alternative to classic tissue biopsy for prevention, diagnosis, and monitoring of patients. Accurate assessment of disease activity could allow measurable gain in the treatment time course. We have developed a biomarker measuring serum levels of matrix metalloproteinase (MMP)-degraded type I collagen (C1M). C1M is destroyed by cathepsin K, therefore making this a marker specific to soft tissue type I collagen. As MMPs are often expressed by inflammatory cells, C1M reflects the remodeling potential of the body and cell inflammation. We have previously seen that C1M is associated with early death in postmenopausal women and with age-related pathologies such as fibrosis, cancer, and rheumatoid arthritis.

Despite the enormous impact dysregulation of type I collagen metabolism has on human health, little is known about the genetic architecture of collagen metabolism. Greater insight into this topic could therefore help with the understanding of age-related pathologies.

In this study, we performed a genome-wide screen for association C1M in a population of 4,981 postmenopausal women from the Prospective Epidemiologic Risk Factor (PERF) study using ca. 7.6 million genetic markers. Our objectives were to comprehensively identify genetic variants influencing the serum levels of C1M and test the relationship between identified genetic variants and common age-related pathologies and finally to investigate the role of collagen metabolism in these.

Methods

Study design

The PERF study was a follow-up study of Danish postmenopausal women aiming at identifying risk factors associated with age-related diseases. A total of 5,855 women were enrolled at baseline (PERF I) between 1999 and 2001. Subjects in PERF have previously participated in clinical randomized placebo-controlled studies or were screened without being randomized for previous studies at the Center for Clinical and Basic Research. We performed a study inclusion process in 3 steps (figure e-1, links.lww.com/NXG/A317): we first collected subjects with demographic, serum, and blood biomarker measurements available with a missingness cutoff of less than 5%. Subjects without genotypes were excluded, and a study population level filter was applied on genotypes to remove cryptic relatedness.

Standard protocol approvals, registrations, and patient consents

The study was conducted in accordance with the International Conference on Harmonization–Guideline for Good Clinical Practice, and the study protocol was approved by the local ethics committees. All participants signed an informed consent, allowing future analysis to be performed.

Baseline measurements and data collection

At baseline, the subjects completed an interview with a doctor or a nurse covering questions related to physical health, demographics, lifestyle, and medical history. Fasting serum and DNA samples were collected from subjects who gave written consent for this specific analysis (n = 5,668 and n = 5,553, respectively).

C1M, caspase-degraded TAU (TAU-C) and ADAM10-degraded TAU (TAU-A) were measured blinded in serum by ELISA in a CAP-certified laboratory as previously described. Lymphocyte and neutrophil counts were determined using an automated blood cell analyzer (Sysmex, Kobe, Japan). Serum cholesterol and triglycerides were...
measured using the Advia 1800 analyzer (Siemens Healthcare Diagnostics, Munich, Germany).

Complete hospital disease history of the subjects was obtained for the period 1974–2014 by linking each individual’s unique personal identification number (CPR number) with the Danish patient registries on December 31, 2014, corresponding to end of study. Subjects of the study were anonymized, and Central Personregister (CPR) numbers were not made available at any point of the study. Patient registry information was available for 5,602 subjects.

**Disease phenotype definition**

Fourteen disease phenotypes have been defined as all-time incidence of an event based on data available from multiple sources: biochemical marker levels, physiologic measurements, all-time incidence hospital records, death registry, questionnaire data from baseline consultation, on previous medical history at PERF I and PERF II enrollment times, and cognitive tests. The detailed list of included phenotypes and their inclusion criteria is provided in table e-2 (links.lww.com/NXG/A317). In this study, dementia was defined as (1) International Classification of Diseases, Tenth Revision (ICD-10) codes F01-F03 G31-G32, (2) The Short Blessed Test ≥10 and the Category Fluency Test ≤14, or (3) dementia noted in questionnaires. Alzheimer disease was defined as (1) ICD-10 codes F00+G30 or (2) Alzheimer disease noted in questionnaires.

**Genotyping and imputation**

Of 5,553 DNA samples collected, 5,516 have been genotyped successfully. Genotyping was performed using a custom-made Illumina Global Screening Array (693,143 probes) in collaboration with deCODE Genetics, Iceland. Single nucleotide polymorphism (SNP) imputation was performed using the Michigan Imputation Server.15 The reference panel used for this step is the HRC r1.1.2016, EUR population. Phasing was performed with ShapeIt2 and the imputation with Minimac3. Using a curated set of 563,532 probes, we imputed 39131581 positions. Positions are reported as in the GRCh37 reference.

**Probe-level filtering**

Standard probe-level filtering has been performed using a minimum probe call rate of 97%, a minor allele frequency (MAF) greater than or equal to 1%, and a Hardy-Weinberg equilibrium p value cutoff greater than or equal to 1e-6. No filtering on multiallelic SNPs has been performed. In total, 534,710 probes and 7,672,338 imputed positions were tested in the association screening, respectively.

**Individual-level filtering**

To address possible cryptic relatedness between subjects, we calculated an identity-by-descent coefficient using Plink16–genome function. Inbreeding coefficient was calculated using the plink –ibc function. We removed patients, on a 1 side of a pair basis, using a minimum PI_HAT cutoff value of 0.1875, and a cutoff of less than ~0.1 or greater than 0.1 was

---

**Table 1** Baseline characteristics of the PERF study

| Variable                        | N   | PERF study group |
|---------------------------------|-----|-----------------|
| Baseline age, y, mean (SD)       | 4,891| 70.1 (6.5)      |
| Education, n                    | 4,887|                 |
| Primary school                  | 3,497|                 |
| High school                     | 1,044|                 |
| University                      | 346  |                 |
| Height, cm, mean (SD)           | 4,891| 161.0 (5.9)     |
| Weight, kg, mean (SD)           | 4,891| 67.7 (11.6)     |
| BMI, kg/m², mean (SD)           | 4,891| 26.1 (4.2)      |
| Systolic, mm Hg, mean (SD)      | 4,889| 150.1 (24.3)    |
| Diastolic, mm Hg, mean (SD)     | 4,891| 81.8 (11.4)     |
| Biochemistry, mean (SD)         |      |                 |
| Total cholesterol: serum, mL    | 4,891| 6.3 (1.1)       |
| Triglycerides: serum, mL        | 4,891| 1.4 (0.6)       |
| LDL cholesterol: serum, ng/mL   | 1,002| 3.5 (1.0)       |
| HDL cholesterol: serum, ng/mL   | 2,896| 1.7 (0.4)       |
| Neutrophils count, 10e9/L       | 4,632| 58.5 (8.9)      |
| Lymphocyte count, 10e9/L        | 4,632| 32.9 (8.1)      |
| C1M: serum, ng/mL               | 4,891| 51.2 (47.9)     |
| TAU-C: serum, ng/mL             | 4,885| 22.0 (12.0)     |
| TAU-A: serum, ng/mL             | 4,883| 28.2 (16.1)     |
| SBT score, n                    | 4,855|                 |
| 0–9                             | 4,684|                 |
| 10 or more                      | 171  |                 |
| CFT score, n                    | 4,851|                 |
| Over 14                         | 4,333|                 |
| 14 or less                      | 518  |                 |
| Smoking, n                      | 4,891|                 |
| Never                           | 2,306|                 |
| Past                            | 1,496|                 |
| Current                         | 1,089|                 |
| Alcohol, n                      | 4,871|                 |
| <7 alcohol units per week       | 3,271|                 |
| >7 alcohol units per week       | 1,600|                 |

Abbreviations: BMI = body mass index; CFT = Category Fluency Test with animal naming; C1M = MMP-degraded type I collagen; MMP = matrix metalloproteinase; PERF = Prospective Epidemiologic Risk Factor; SBT = Short Blessed Test; TAU-A = ADAM10-degraded TAU; TAU-C = caspase-degraded TAU.

Characteristics of the subjects included in the study are shown (n = 4,981).
applied to the $F_{hat2}$ coefficient. In total, 136 of 5,516 individuals have been removed in this step.

**Principal component analysis**

Population-based genetic variation in the data set was captured using EIGENSTRAT Smartpca 7.2.0\textsuperscript{17,18} to perform an iterative principal component analysis (PCA) of the study population with available genotypes ($n = 5,106$) on the nonimputed filtered variants using the default parameters. The scree plot of the first 10 components and the PCA plot of the 2 leading components are shown in figures e-6 and e-7 (links.lww.com/NXG/A317). The first 3 components capture the largest share of explained variance (0.3%).

**Linear regression**

Linear additive regression was performed on the genome-wide association study (GWAS) population ($n = 4,981$) to identify genetic associations with log2-transformed serum C1M levels using plink v1.90p adjusted for baseline age, body mass index (BMI), and the 3 leading principal components, and prebaseline occurrence of cancer, inflammatory arthropathies, and spondylopathies, defined as events up to 1 year after baseline to address potential influence on C1M levels. The plink switches –allow-no-sex and –keep-allele-order were also used. Subjects with missing biomarker levels were not included in the analysis. Conservative significance thresholds based on the number of screened variants were defined as equal to 6.5e-9, i.e., 0.05/N, and 1.3e-7, i.e., 1/N, $N = 7672,338$, for the genome-wide and suggestive values, respectively. The Manhattan plot visualization has been made using the R package qqman\textsuperscript{19} and color palette from ggsci.

**Phenotype-genotype association analysis**

Phenotype-genotype association analysis (PheWAS) was performed using a logistic regression model, corrected for baseline age, BMI, and the 3 leading principal components. The $Z$-statistic of the regression was reported and visualized with a heatmap, generated by the R package ComplexHeatmap.\textsuperscript{20}

**Biomarker boxplot visualization**

Serum and blood biomarker level distributions across time to occurrence of Alzheimer disease relative to baseline were shown using R ggplot2 functions. Tests of biomarker distributions between time intervals were performed using an analysis of variance on log-transformed values with post hoc comparisons with the Holm-Sidak multiple comparisons test.

**Data availability**

The original data of the PERF study and the linkage data from various health registries are currently stored at Nordic Bioscience. Access to this database will be granted, on condition that researchers have appropriate ethical permission and sign the appropriate Material Transfer Agreement form.
Results

Study design
A total of 5,855 subjects have been included in the PERF study at baseline. For a majority of these, genotyping (n = 5,516) has been performed, and serum and blood biomarkers (n = 5,668) were measured. More details are available in the Material and Methods section and in the cohort profile presentation paper.12 A GWAS was performed on the maximum population with both genotypes, demographic and biomarker measurements, and data availability from the Danish patient registries (n = 4,891). Table 1 summarizes the baseline characteristics of the study group.

Genetic associations with C1M levels
We identified 22 genome-wide significant SNPs associated with C1M, located primarily in the chr19q13.32 APOE-C1/TOMM40 gene locus (figure 1 and table e-1, links.lww.com/NXG/A317). Within this cluster, the most significant association with APOE variants was rs429358—located in exon 4, and which combination with rs7412 is commonly known as APOE e4 (p = 1.38e-17, effect size: −0.162; 95% confidence interval [CI] −0.199 to −0.1248 log2 ng/mL per additional minor allele) and rs769449 (p = 7.15e-16, effect size: −0.166; 95% CI −0.209 to −0.129 log2 ng/mL). Nine genetic variants near the APOCI gene were significantly associated with C1M including rs4420638 (p = 8.44e-14) and 4 associations with PVRIL2 variants including rs6857 (p = 2.07e-13). SNPs in this locus are relatively common (MAF from 0.13 to 0.25) and in moderate linkage disequilibrium, see figures e-2 and e-4 for the LD structure and the locus plot of the region. An additional 6 suggestive associations were found in this locus, centered around TOMM40. All genome-wide significant associations in chr19 had a negative effect size (β-coefficients ranging −0.108 to −0.169 log2 ng/mL). We identified 10 suggestive associations near the gene C-reactive protein (CRP), mostly located downstream (locus plot shown in figure e-3). The nearest associated variant was rs3093068 (p = 1.09e-7) with a positive effect (β = 0.172; 95% CI 0.108–0.235 log2 ng/mL).

Phenotype-genotype association analysis
A PheWAS was performed on the 38 SNPs screened by the GWAS (figure 2 and table e-5, links.lww.com/NXG/A317). The variants identified in APOE-C1/TOMM40 cluster were positively associated with the incidence of dementia and Alzheimer disease (Z-statistic range 4.89–9.47) as well as...
hyperlipidemia ($Z$-statistic range 3.06–6.32). It was previously shown that rs429358 was associated with the increased serum total cholesterol level and depression severity, and the authors hypothesized $APOE$ variants to be detrimental to recovery of nerve function after stroke. By performing a GWAS, corrected for BMI, baseline age, and population structure, we confirmed that $APOE$ variants were associated with total cholesterol, with top hit rs7412, $p = 3.4e^{-47}$ (figure e-5). Functional analyses (table e-6) showed an enrichment of GO terms related to regulation of cholesterol biosynthetic process, lipoprotein particle receptor binding, protein-lipid complex binding, and low-density lipoprotein particle remodeling. In addition, these variants were associated with decreased incidence of neoplasms ($Z$-statistic range $-3.04$ to $-2.12$).

Variants located in the chr1 $CRP$ gene locus had a weak association with cardiovascular traits ($Z$-statistic range 1.33–1.67) and a negative association with dementia and Alzheimer disease. 

**C1M degradation in preclinical dementia/Alzheimer disease**

We have previously shown that biomarkers of ADAM10- and caspase-degraded tau (TAU-A and TAU-C) were negatively associated with dementia in PERF. In the current study, we noted that there was a strong correlation between C1M and tau degradation biomarker levels ($r = 0.54$ between TAU-C and C1M, $r = 0.58$ between TAU-A and C1M, figure 3, A and B). By grouping dementia/Alzheimer disease incidences into 2 bins: (1) dementia/Alzheimer disease at baseline (diagnosed less than 2 years after baseline) and (2) preclinical dementia/Alzheimer disease (diagnosed more than 2 years after baseline), and looking at C1M levels, we observed that women with preclinical dementia/Alzheimer disease had lower levels of C1M compared with women with dementia/Alzheimer disease at baseline and women with no cognitive impairment (never diagnosed with dementia/Alzheimer disease, figure 3C). We also found that women diagnosed with both dementia/Alzheimer disease at baseline and preclinical
Genome-wide association analysis of C1M levels in our study population identified two main genomic regions: chr19 q13.32 APOE-C1/TOMM40 cluster and chr1q23.3, which encompasses the gene CRP. SNPs in the APOE cluster were dominantly associated with lower C1M levels, whereas SNPs within CRP were associated with a higher C1M level. Power calculations performed using a type I error cutoff at 5e-8 showed that the association found in the APOE locus was sufficiently supported (statistical power for rs429358: 0.68) while findings within CRP would require a larger sample size.

Although our PheWAS analysis did not identify strong associations with SNPs within the CRP gene locus, except for a few markers previously associated with heart disease, we found that the variants associated with C1M in the APOE-C1/TOMM40 locus were strongly associated with dementia and Alzheimer disease, as for example, the well-studied variant rs429358, which is a risk factor for neural regeneration in late-onset Alzheimer disease.23–25

The analysis of APOE ε4 genotype frequencies (rs429358 and rs7412) in our study population (tables e-3 and e-4, links.lww.com/NXG/A317) showed that they were in agreement with those observed in the Danish general population26 and that APOE ε4 genotypes were overrepresented in the individuals with Alzheimer disease ($\chi^2$ test $p = 3.6e-16$). To further study the link between type I collagen remodeling and in preclinical dementia/Alzheimer disease, we showed that C1M levels could be stratified by their APOE ε4 genotypes, with significantly lower C1M levels in APOE ε4 double carriers compared with subjects without ε4 alleles.

When we looked further into the biomarker dynamics in subjects with cognitive impairment, we found that C1M correlated with tau degradation markers (TAU-A and TAU-C). We have previously seen that high levels of the tau degradation markers were associated with a lower risk of preclinical dementia and Alzheimer disease.22 A plausible explanation for this could be linked to microglial activation. In early stages of dementia and Alzheimer disease, microglial activation is believed to be neuroprotective by enhancing phagocytosis and degradation of β-amyloid and tau,27,28 a process that may result in less release of tau degradation products to the periphery. In later stages, where microglia become overactivated, they lose their phagocytic abilities, resulting in uncontrolled inflammation releasing degraded tau to the periphery.29

In this study, C1M levels were also lower in subjects with preclinical dementia/Alzheimer disease compared with subjects diagnosed close to baseline and subjects with no cognitive impairment. We hypothesize that APOE ε4 carriers are born with a low remodeling potential (low C1M levels), which increases the risk of dementia/Alzheimer disease along with other age-related diseases because of inefficient clearance of damaged proteins and cells (figure 5B). In subjects with low
dementia/Alzheimer disease had lower levels of neutrophils and higher levels of lymphocytes compared with women with no cognitive impairment (figure 3, D and E).

**C1M degradation stratified by APOE ε4 genotype in preclinical dementia/Alzheimer disease**

We looked at the C1M degradation biomarker levels in the subpopulation of women with preclinical dementia/Alzheimer disease ($n = 370$), stratified by their APOE ε4 genotypes, obtained from the 2 variants rs429358 and rs7412 (figure 4). We see an allele dose effect between women with no APOE ε4 allele and women carrying 1 and 2 APOE ε4 alleles (1-sided Wilcoxon test $p < 0.0001$). This result, in line with the GWAS, further demonstrates the negative association between APOE genotypes and C1M in preclinical dementia/Alzheimer disease.

**Discussion**

We investigated the genetic component in the variation of C1M in a population of postmenopausal Danish women. This biomarker reflecting inflammation and remodeling potential of the body has been previously described to be associated with age-related diseases such as cancer, fibrosis, and rheumatoid arthritis.8–11 The aim of our study was to systematically discover genetic factors of variation of C1M and link these variants to age-related diseases. The PERF cohort offered an ideal exploratory environment combining single nucleotide variants, demographic, and electronic hospital care history.
remodeling potential, the clearance of amyloid-β will be inefficient, leading to formation of amyloid-β oligomers and plaques. This will in turn initiate chronic inflammatory processes and thereby increase C1M levels in the later stages of the disease (figure 5A).

We also found that women with either dementia/Alzheimer disease baseline or preclinical dementia/Alzheimer disease had increased neutrophil count and decreased lymphocyte count compared with subjects with no cognitive impairment. It is well known that chronic inflammation worsen during the course of dementia and Alzheimer disease. 2,30–32 Proinflammatory cytokines have been detected in the periphery, indicating that a strong innate immune response is occurring throughout disease progression and triggered by the dysregulation of the Aβ peptide. 33 A decreased lymphocyte count have on the other hand been shown lead to greater accumulation of amyloid Aβ plaques and microglia activation in mice, 34 indicating that the adaptive immune system is dysfunctional in subjects with Alzheimer disease.

Together, the results of this study suggest that targeting inflammation or the remodeling potential could be relevant for therapeutic interventions and preventive strategies; however, the study population should be stratified according to parameters including APOE genotype and inflammatory phenotype.

There are a few potential limitations in our study. Type I collagen is ubiquitous and may be affected by external factors. We addressed this by correcting for disease conditions previously associated with elevated C1M levels. Medications could also potentially modulate type I collagen remodeling. Use of medication was available as survey information for a part of the follow-up group (n = 1,856). Sensitivity analysis performed for frequently reported medications, e.g., estradiol and paracetamol, showed that the effect sizes of the reported SNPs did not deviate more than 1.5%, and therefore, medication could be disregarded from the GWAS model. Neuropathologic confirmation was not available for this study. Diagnosis was based on a combination of data from the Danish national patient registry, questionnaires, and cognitive tests (table e-2, links.lww.com/NXG/A317). Finally, gene expression was not collected in the PERF study. Evaluating the impact of genetic variants on gene expression could contribute to better understand the impact of type I collagen remodeling in cognitive disorders.

In conclusion, our study uncovers the link of type I collagen remodeling in lipoprotein balance and onset of dementia and Alzheimer disease. Our results suggest that blood-based measurements of inflammation and tissue remodeling could be relevant as a first-line therapeutic intervention.
Acknowledgment
The authors thank Unnur Stýrkarsdóttir for her valuable help in the genotyping of the PERF samples and insightful discussions; Line Mærsk Staunstrup for her contribution to the PERF database curation; and Peder Frederiksen for valuable scientific discussion.

Study funding
J.P.M. Blair was funded by a PhD grant from the Danish Research Foundation.

Disclosure
M.-H.E. Tang and C.L. Bager are employed at ProScion. A.-C. Bay-Jensen, K. Henriksen, and M.A. Karsdal are employed at Nordic Bioscience. A.-C. Bay-Jensen, C. Christiansen, and M.A. Karsdal hold stocks in Nordic Bioscience. No other potential conflicts of interest relevant to this article were reported. Go to Neurology.org/NG for full disclosures.

Publication history
Received by Neurology: Genetics May 13, 2020. Accepted in final form July 28, 2020.

Appendix
Authors

| Name                  | Location          | Contribution                              |
|-----------------------|-------------------|-------------------------------------------|
| Man-Hung Eric Tang, PhD | ProScion, Herlev, Denmark | Designed and conceptualized the study; analyzed and interpreted the data; and drafted the manuscript for intellectual content |
| Joseph P.M. Blair, MSc | ProScion, Herlev, Denmark | Analyzed and interpreted the data and revised the manuscript for intellectual content |
| Cecille Liv Bager, PhD | ProScion, Herlev, Denmark | Major role in the acquisition of data; interpreted the data; and drafted the manuscript for intellectual content |
| Anne-Christine Bay-Jensen, PhD | Nordic Bioscience, Herlev, Denmark | Interpreted the data and revised the manuscript for intellectual content |
| Kim Henriksen, PhD | Nordic Bioscience, Herlev, Denmark | Interpreted the data and revised the manuscript for intellectual content |
| Claus Christiansen, MD | Nordic Bioscience, Herlev, Denmark | Major role in the acquisition of data and revised the manuscript for intellectual content |
| Morten Asger Karsdal, PhD | Nordic Bioscience, Herlev, Denmark | Interpreted the data and drafted the manuscript for intellectual content |

References
1. Gubbels Bupp MR. Sex, the aging immune system, and chronic disease. Cell Immunol 2015;294:102–110.
2. Mishra A, Britton RD. Inflammation: bridging age, menopause and APOEε4 genotype to Alzheimer’s disease. Front Aging Neurosci 2018;10:312.
3. Wynn TA. Common and unique mechanisms regulate fibrosis in various fibroproliferative diseases. J Clin Invest 2007;117:524–529.
4. Karsdal MA, Nielsen SH, Leeming DJ, et al. The good and the bad collagens of fibrosis: their role in signaling and organ function. Adv Drug Deliv Rev 2017;121:43–56.
5. Leeming DJ, Bay-Jensen AC, Yasuda E, Larsen MR, Henriksen K, Karsdal MA. Post-translational modifications of the extracellular matrix are key events in cancer progression: opportunities for biochemical marker development. Biomarkers 2011;16:193–205.
6. Genovese F, Karsdal MA. Protein degradation fragments as diagnostic and prognostic biomarkers of connective tissue diseases: understanding the extracellular matrix message and implication for current and future serological biomarkers. Expert Rev Proteomics 2016;13:213–225.
7. Leeming D, He Y, Vidal-S, et al. A novel marker for assessment of liver matrix remodeling: an enzyme-linked immunosorbent assay (ELISA) detecting a MMP-generated type I collagen neo-epitope (C1M). Biomarkers 2011;16:616–628.
8. Dragshbek K, Neergaard JJS, Hansen HHB, et al. Matrix metalloproteinase mediated type I collagen degradation: an independent risk factor for mortality in women. Elsino Medicine 2015;2:723–729.
9. Bager CL, Willumsen N, Kehlet SN, et al. Remodeling of the tumor microenvironment predicts increased risk of cancer in postmenopausal women: the Prospective Epidemiologic Risk Factor (PERF) I study. Cancer Epidemiol Biomarkers Prev 2016;25:1348–1355.
10. Siebahr A, Bay-Jensen AC, Leeming DJ, et al. Serological identification of fast progressors of structural damage with rheumatoid arthritis. Arthritis Res Ther 2013;15:R86.
11. Jenkins KG, Simpson JK, Saini G, et al. Longitudinal change in collagen degradation biomarkers in idiopathic pulmonary fibrosis: an analysis from the prospective, multicentre PROFILE study. Lancet Respir Med 2015;3:462–472.
12. Neergaard JS, Dragshbek K, Kehlet SN, et al. Cohort profile: the Prospective Epidemiologic Risk Factor (PERF) study. Int J Epidemiol 2017;46:1104–1104.
13. Henriksen K, Wang Y, Sørensen MG, et al. An enzyme-generated fragment of tau measured in serum shows an inverse correlation to cognitive function. PLoS One 2013;8:e64990.
14. Henriksen K, Byrjalsen I, Christiansen C, Karsdal MA. Relationship between serum levels of tau fragments and clinical progression of Alzheimer’s disease. J Alzheimers Dis 2015;53:1331–1341.
15. Das S, Forer L, Schionherr S, et al. Next-generation genotype imputation service and methods. Nat Genet 2016;48:1284–1287.
16. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. GigaScience 2015;4:7.
17. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet 2006;38:904–909.
18. Patterson N, Price AL, Reich D. Population structure and eigenanalysis. PLoS Genet 2006;2:e190. Available at: netlib.org/lapack. Accessed November 15, 2019.
19. Turner SD. qpman: an R package for visualizing GWAS results using Q-Q and Manhattan plots. bioRxiv 2014 May 14. doi:10.1101/005165.
20. Gu Z, Ellis R, Schlessner M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics 2016;32:2847–2849. Available at: bioconductor.org/packages/devel/bioc/html/ComplexHeatmap.html. Accessed November 15, 2019.
21. Li XB, Wang J, Xu AD, et al. Apolipoprotein E polymorphisms increase the risk of post-stroke depression. Neural Regen Res 2016;11:1790–1796.
22. Neergaard JS, Dragshbek K, Christiansen C, Karsdal MA, Brin S, Henriksen K. Two novel blood-based biomarker candidates measuring degradation of tau are associated with dementia: a prospective study. PLoS One 2018;13:e0194802.
23. Guerreiro R, Ross OA, Kun-Rodrigues C, et al. Investigating the genetic architecture of dementia with Lewy bodies: a two-stage genome-wide association study. Lancet Neurol 2018;17:64–74.
24. Cosentino S, Scarmetrics N, Helmer R, et al. APOE ε4 allele predicts faster cognitive decline in mild Alzheimer’s disease. Neurology 2008;70:1842–1849.
25. Yi L, Wu T, Luo W, Zhou W, Wu J. A non-invasive, rapid method to genotype late-onset Alzheimer’s disease-related apolipoprotein E gene polymorphisms. Neurology Regen Res 2014;9:69–75.
26. Rasmussen KL, Tylbyaerg Hansen A, Nordestgaard BG, Frikke-Schmidt R. Plasma levels of apolipoprotein E and risk of dementia in the general population. Ann Neurol 2015;77:301–311.
27. Frautschy SA, Yang F, Irizarry M, et al. Microglial response to amyloid plaques in APPsw transgenic mice. Am J Pathol 1998;152:307–317.
28. Luo W, Liu W, Hu X, Hanna M, Caravaca A, Paul SM. Microglial internalization and degradation of pathological tau is enhanced by an anti-tau monoclonal antibody. Sci Rep 2015;5:11161.
29. Mrak RE, Griffin WST. Glia and their cytokines in progression of neurodegeneration. Neurobiol Aging 2005;26:349–354.
30. Mueller MM, Fusenig NE. Friends or foes: bipolar effects of the tumour stroma in cancer. Nat Rev Cancer 2004;4:339–349.
31. Rosenberg GA. Matrix metalloproteinases and their multiple roles in neurodegenerative diseases. Lancet Neurol 2009;8:205–216.
32. Bruno MA, Mufson EJ, Wuu J, Cuello AC. Increased matrix metalloproteinase 9 activity in mild cognitive impairment. J Neuropathol Exp Neurol 2009;68:1309–1318.
33. Wang WT, Tan MS, Yu JT, Tan L. Role of pro-inflammatory cytokines released from microglia in Alzheimer’s disease. Ann Transl Med 2015;3:136.
34. Marsh SE, Abud EM, Lakatos A, et al. The adaptive immune system restrains Alzheimer’s disease pathogenesis by modulating microglial function. Proc Natl Acad Sci USA 2016;113:E1316–E1325.