Urinary proteomic signature of exercise intolerance in patients at risk of heart failure: evidence from the HOMAGE trial

Chu-Hao Liu,[a,b,q] Dries S. Martens,[c,q] De-Wei An,[a,q] Justyna Siwy,[d] Agnieszka Latosinska,[d] Pierpaolo Pellicori,[e] Job A. J. Verdonschot,[f] Fozia Z. Ahmed,[g] Fang-Fei Wei,[h] Patrick Rossignol,[i] Johannes Petutschnigg,[j] Stephane Heymans,[f] Joe J. Cuthbert,[k] Yu-Ling Yu,[b,l] Nicolas Girerd,[i] Andrew L. Clark,[k] Peter Verhamme,[m] Dong-Yan Zhang,[a] Yan Li,[a] Tim S. Nawrot,[c,n] John G. Cleland,[e] Faiez Zannad,[i] Harald Mischak,[d] and Jan A. Staessen,[a,b,o,p,*] on behalf of the HOMAGE investigators[r]

[a]Department of Cardiovascular Medicine, Shanghai Key Laboratory of Hypertension, Shanghai Institute of Hypertension, State Key Laboratory of Medical Genomics, National Research Centre for Translational Medicine, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China
[b]Non-Profit Research Association Alliance for the Promotion of Preventive Medicine (APPREMED), Mechelen, Belgium
[c]Centre for Environmental Sciences, Hasselt University, Hasselt, Belgium
[d]Mosaiques-Diagnostics GmbH, Hannover, Germany
[e]School of Cardiovascular and Metabolic Health, University of Glasgow, Glasgow, UK
[f]Department of Cardiology, Maastricht University Medical Centre, Maastricht, the Netherlands
[g]Division of Cardiovascular Sciences, School of Medical Sciences, Faculty of Biology, Medicine and Health, Manchester Academic Health Science Centre, University of Manchester, Oxford Road, Manchester, UK
[h]Department of Cardiology, First Affiliated Hospital of Sun Yat-Sen University, Guangzhou, China
[i]Université de Lorraine, Inserm, Centre d'Investigation Clinique Plurithématique 1433, U1116, CHRU Nancy, F-CRIN INI-CRCT, Nancy, France
[j]Department of Internal Medicine and Cardiology, Campus Virchow Klinikum, Charité University Medicine Berlin, Berlin Institute of Health and German Centre for Cardiovascular Research, Berlin, Germany
[k]Department of Cardiology, University of Hull, Castle Hill Hospital, East Riding of Yorkshire, Cottingham, UK
[l]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden
[m]Department of Cardiovascular Sciences, Centre for Molecular and Vascular Biology, University of Leuven, Leuven, Belgium
[n]Research Unit Environment and Health, Department of Public Health and Primary Care, University of Leuven, Leuven, Belgium
[o]First Department of Cardiology, Interventional Electrocardiology and Hypertension, Jagiellonian University Medical College, Kraków, Poland
[p]Biomedical Science Group, University of Leuven, Leuven, Belgium

eBioMedicine
2026;130: 106406
Published Online xxx
https://doi.org/10.1016/j.ebiom.2026.106406

Summary

Background Urinary proteomic profiling (UPP) provides insights in disease mechanisms and origin of symptoms. Using UPP, this study aimed at deepening insight in the biology of exercise tolerance.

Methods In the HOMAGE trial, 268 patients at risk of heart failure underwent the incremental shuttle walk test (SWT) and UPP by capillary electrophoresis coupled with mass spectrometry at baseline (discovery) and the 9-month final visit (replication). Sequencing of 1498 urinary peptides identified 170 non-collagen and 40 collagen-derived proteins. Ten exercise-related variables, including heart rate and blood pressure responses, symptoms and walking distance were summarised into a single factor, higher values indicating exercise intolerance. In exploratory analyses, exercise intolerance was related to the UPP, first in linear and logistic regression models, considering one peptide at a time, and next by elastic net regression considering the peptides retained in the previous step. Replicated proteins were subjected to pathway analysis.

Findings Twenty-nine non-collagen and 31 collagen-derived peptides were associated with reduced exercise capacity with correction for multiple testing. In elastic net regression, exercise intolerance was in >50% of 1000 bootstrap runs associated with the non-collagen proteins FXYD2, GAPDH, HBB, KCNB1, MB, MYOCD, SECTM1, SNX9, TACC3, TMSB4X, and TTN and with collagens COL5A1, COL6A1, COL11A2, and COL28A1. Enriched pathways

*Corresponding author. Alliance for the Promotion of Preventive Medicine, Leopoldstraat 59, Mechelen BE-2800, Belgium.
E-mail address: jan.staessen@appremed.org (J.A. Staessen).
@jasta49 (J.A. Staessen).
[q]Joint first authors who contributed equally.
[r]The HOMAGE Investigators are listed in the Appendix p 2.

involved oxygen homoeostasis, oxidative stress regulation, metabolic and developmental processes, and muscle biology.

Interpretation In HOMAGE patients, UPP identified parent proteins regulating exercise endurance, which are involved in oxygenation, vascular and muscle structure and function, maintenance of the circulating volume, and protection against oxidative stress.

Funding European Union.

Keywords: Exercise; Heart failure; Incremental shuttle walk test; Systems biology; Urinary proteomics

Research in context

Evidence before this study

PubMed searches for studies published in English were updated until 17 June 2026, using as search terms "exercise" AND "endurance" AND "proteomics" in the title or abstract. The searches returned 70 articles published from 2008 to 2026. All summaries were examined for relevance. The full text of 13 articles was reviewed to establish the current state of knowledge. Tissue and circulating proteomics, without or with genetic profiling and multi-omics, were applied in experimental, observational, and interventional studies to assess muscle physiology, exercise endurance and trainability of healthy individuals, athletes, or patients with various disorders. However, no study related exercise tolerance to the urinary proteome in spite of the molecular and pathophysiological information that can be derived from sequenced urinary peptides that identify parent proteins.

Added value of this study

This study summarised 10 variables measured during an incremental shuttle walk test into a single factor. The analysis showed a relation between exercise tolerance and the parent proteins of urinary peptides, which included FXYD2, GAPDH, HBB, KCNB1, MB, MYOCD, SECTM1, SNX9, TACC3, TMSB4X, TTN, COL5A1, COL6A1, COL11A2, and COL28A1. These proteins are involved in pathways related to oxygen homoeostasis, oxidative stress regulation, metabolic and developmental processes, and cytoskeletal organisation of muscle cells.

Implications of all the available evidence

The current study proved the feasibility of using a single factor combining 10 exercise-related features and showed correlations with the urinary proteome of patients at high risk of heart failure. Modifying the biological pathways involved in exercise limitation may be a strategy for future therapy.

Introduction

Exercise capacity provides important diagnostic and prognostic information in patients with cardiac[1,2] or pulmonary[3] conditions. Incremental cardiopulmonary exercise testing with metabolic gas exchange measurement (CPET) is the standard method to assess maximal aerobic capacity.[4] However, CPET is a complex procedure. The incremental shuttle walk test (SWT) is a symptom-limited maximal exercise test (Appendix, pp 3–7) used to assess exercise tolerance in patients with chronic heart failure (HF).[5] The correlation coefficients between the SWT distance covered and peak oxygen consumption measured by CPET range from 0.67 to 0.95.[6] Reduced exercise endurance, as reflected by deficient chronotropic and haemodynamic responses to exercise, is an independent predictor of cardiovascular endpoints, post-surgical complications, and all-cause mortality, especially in older adults, patients with respiratory disease, a history of myocardial infarction, cancer, or chronic kidney disease.[7,8] These observations justify further research in exercise intolerance for the timely diagnosis of subclinical changes in left ventricular structure and function, in particular HF with preserved ejection fraction, which is prevalent among women and is often only characterised by unexplained breathlessness.[9]

Sequenced peptides in the urinary proteome identify parent proteins.[10,11] Extensive PubMed searches (updated until 17 June 2026) and comprehensive reviews[12,13] revealed that proteomics is widely used to assess muscle physiology,[14–17] exercise endurance, and trainability in healthy individuals,[18] athletes,[19–22] and patients with metabolic disorders.[23–25] However, there were no reports of the relation between exercise tolerance and the parent proteins identified by urinary proteomic profiling (UPP) with the exception of one acute intervention study with little relevance in 10 firefighters.[26] This knowledge gap was addressed in the Heart Omics in Ageing (HOMAGE) study,[27] an open-label, randomised clinical trial, in which patients at risk of

HF were randomised to usual therapy with or without spironolactone. Of 527 randomised patients, 268 patients underwent SWT testing and had the UPP measured at baseline and at the 9-month terminating visit (Fig. 1).

Methods

Study participants

HOMAGE is a multicentre open-label trial with blinded endpoint evaluation (registration number: NCT02556450).[27] The protocol was approved by the Greater Manchester Central Research Ethics Committee (reference: 16/NW/0012; EudraCT number: 2015-000413-48) and by each clinical site's local Ethics Committee (Appendix p 11). All patients signed an informed written consent form with no provisions for data sharing, given the European General Data Protection Regulation.[28]

Patients of either sex (self-reported), aged ≥60 years, were eligible, provided that they were at high risk of developing HF, because they already had or were likely to develop coronary heart disease (CHD). Patients had serum N-terminal pro-B-type natriuretic peptide (NT-proBNP) of 17–140 pmol/L (125–1000 ng/L) or serum B-type natriuretic peptide (BNP) of 10–81 pmol/L (35–280 ng/L). These ranges excluded patients at low HF risk and those with advanced disease requiring further investigation and intensive treatment. Patients with acute or chronic HF were therefore excluded from the trial. Of 877 screened patients, 527 were randomised to spironolactone 25–50 mg per day (n = 265) in

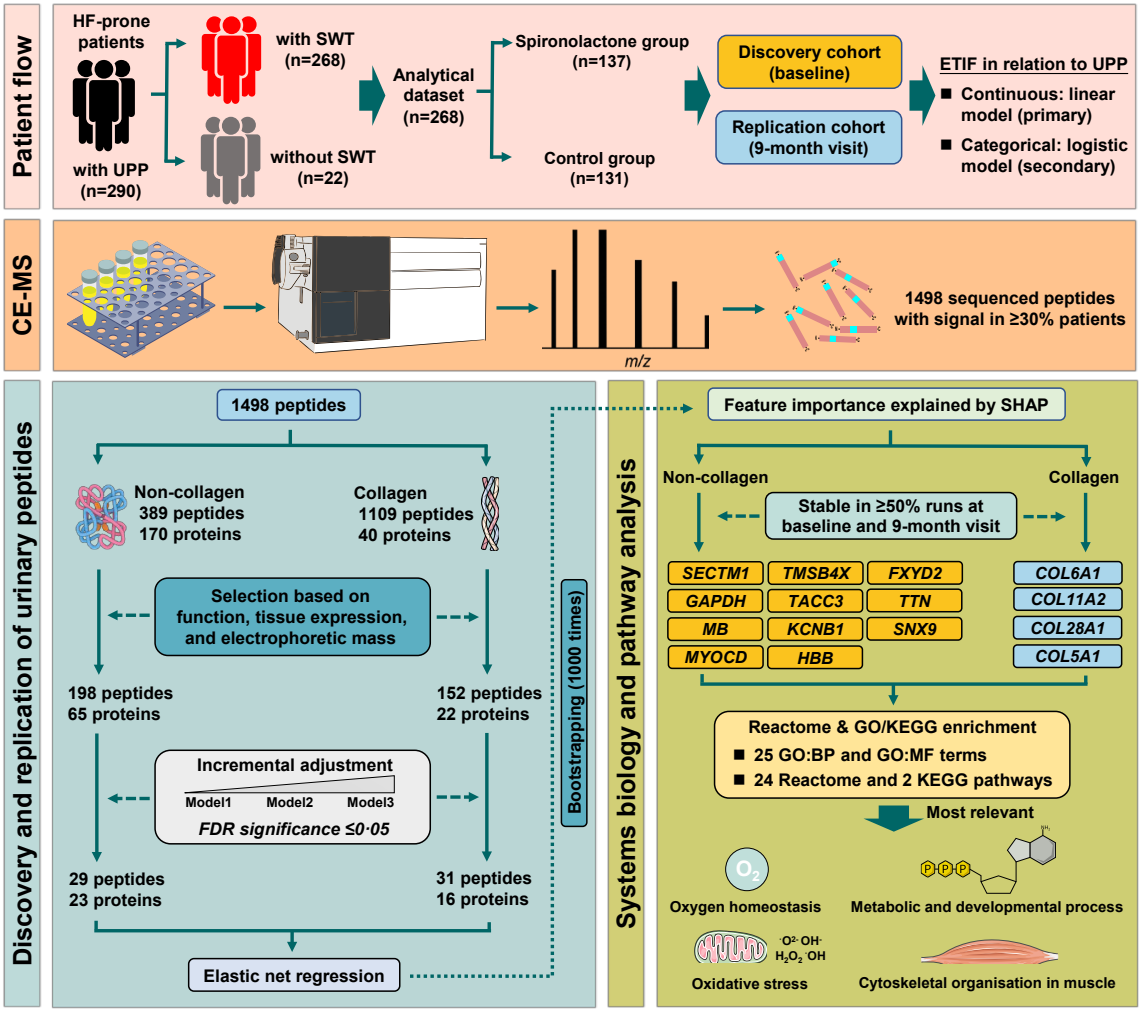

Fig. 1: Diagram illustrating the selection of patients, the technique of capillary electrophoresis coupled with mass-spectrometry, the selection of sequenced peptides identifying the parent proteins, and the flow of the analyses (exploratory analyses using generalised linear models or logistic models with three steps of increasingly stringent adjustment, the main analysis based on elastic net regression, feature importance explained by SHAP, and the pathway analysis).

addition to usual treatment or usual treatment alone (n = 262).[27] Of the UPP data, only sequenced urinary peptides with a detectable signal in at least 30% of patients at baseline and the 9-month terminating visit (n = 290) were analysed. Of these 290 patients, 22 did not undergo SWT testing, leaving 131 patients from the control group and 137 randomised to spironolactone. Previous HOMAGE publications showed that spironolactone had no effect on exercise endurance,[7,8] so that for the current analysis, the 268 patients were pooled. All required exercise data were available (Appendix Fig. S1), with the baseline and 9-month data constituting the discovery and replication datasets. A previous HOMAGE publication on the effects of spironolactone on the UPP[29] suggested that a sample size of 268 patients would be sufficient to address the objectives of the current study. Left ventricular mass was determined in all patients by transthoracic echocardiography during screening and the ejection fraction in 224 patients at baseline and in 180 at the 9-month visit.

Urinary proteomics and biomarkers

Mosaiques-Diagnostics GmbH, Hannover, Germany (MOS), performed UPP profiling for all patients. The methods for sample preparation, capillary electrophoresis coupled with mass spectrometry (CE-MS), peptide sequencing, and for the evaluation, calibration, and quality control of the mass spectrometric data have been previously described in multiple publications[10,11] and in the Appendix, pp 8–10. In the CE-MS step, 29 endogenous urinary peptides were used as internal standards to calibrate signal intensity. This procedure is highly reproducible and addresses, in a single calibration step, both analytical and dilutional variances, including those caused by variability in renal function.[30] The UPP dataset comprised 1498 sequenced urinary peptides with detectable signals in at least 30% of patients. Undetectable peptides were set to the distribution minimum, multiplied by a factor randomly varying from 0.95 to 1.05, to avoid the accrual of an excessive number of peptide levels with the same value (the minimum).[31] The distributions of the urinary peptides (Appendix Fig. S2) were rank normalised by sorting measurements from the smallest to the highest value and then applying the inverse cumulative normal function.[32] Rank-normalised variables have a mean of 0 and a SD of 1. Centring the urinary peptides on the mean also addresses batch effects, such as for instance, those related to the background medical treatment, changes in renal function, or small procedural variability across successive UPP runs.[33]

The UPP dataset consisted of 389 peptides from 170 non-collagen parent proteins and 1109 fragments derived from 40 distinct collagen alpha chains. The number of urinary peptides by which a parent protein is represented increases the likelihood of finding an association with the outcome variable, i.e., exercise tolerance. For instance, 421 peptide fragments originated from COL1A1 (collagen alpha-1[I] chain) and 27 from PIGR (polymeric immunoglobulin receptor). Furthermore, along the renal tubules and along the upper and lower urinary tract, proteases cleave peptides, so that longer peptides represent a greater and therefore more representative fraction of the parent protein.[34] To address both issues, a two-pronged approach was implemented. First, the Protein Atlas (https://proteinatlas.org) and several publications[12–25] were searched for non-collagen proteins and for collagen-alpha chains with functional and expression data in the cardiovascular system, adrenal glands, kidney, skeletal and vascular muscle, and myocardium. The Protein Atlas searches were carefully documented in an Excel report, which is available online as supporting information (Selection of Proteins.xlsx). This file contains the protein symbol, the laboratory identification number, the protein function as described in the Protein Atlas, the expression profile either as RNA or protein in the tissues of interest, the intracellular location of the proteins, a colour code representing the selection, and a justification of the selection. Second, if multiple peptides identified the same parent protein, those with low mass were excluded. This reduced the number of peptide fragments included in statistical analysis to 198 for non-collagen proteins and to 152 for the collagen alpha-chains. Throughout this article, urinary peptides are labelled by the laboratory identification number (Mosaiques-Diagnostics) and parent proteins by their symbol and accession number in the UniProt database (www.uniprot.org). Glomerular filtration rate was estimated (eGFR) from serum creatinine by the Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI) formula.[35] Serum NT-proBNP was measured by electro-chemi-luminescence (ELECSYS® 2010 analyser; Roche Diagnostics, Mannheim, Germany).

Exercise tolerance index

The raw SWT distance (number of shuttles completed) is commonly used by clinicians as an easily implementable substitute for CPET, which requires bicycle testing and metabolic gas exchange measurement in a specialised laboratory.[1–3] However, the SWT distance only captures a single dimension of exercise capacity. Exercise elicits profound haemodynamic responses, including steep increases in blood pressure (BP) and heart rate (Appendix Fig. S3), followed by a partial return to the pre-exercise values.[7,8] The SWT distance provides no information on these haemodynamic responses to exercise, symptoms occurring during the test, such as breathlessness, and on whether the SWT test was stopped on the initiative of the patient (symptom-limited) or the observer (medical indication). All

these factors, 10 in total, were captured by a single variable, of which the derivation is described in the statistical analysis section.

Statistical analysis

For database management and statistical analysis, SAS software, version 9.4, maintenance level 7, was used (SAS Institute Inc., Cary, NC). Deviation from the normal distribution was assessed by the Shapiro–Wilk statistic. For normally distributed data, the central tendency (spread) of the distributions was presented by the arithmetic mean (SD) and for transformed data by the geometric mean or median (interquartile range [IQR]). Arithmetic and geometric means were compared using the large-sample Z-test and proportions by the χ^2 statistic or Fisher exact test, as appropriate based on cell frequencies. Among the clinical and biochemical variables used in the analyses, 26 missing values were imputed, using linear regression: 11 for the waist-to-hip ratio (WHR), 8 for eGFR, and 7 for echocardiographic left ventricular hypertrophy (LVH) as derived from left ventricular mass, indexed to body surface area. Pearson correlation coefficients were used to quantify the interrelations between the urinary peptides, separately for the non-collagen fragments and collagen-derived peptides. Statistical tests were two-sided. The flow of the analyses comprised four steps.

Derivation of the exercise tolerance index

Using data obtained at screening, exercise performance was summarised into a single exercise tolerance index (ETIF), using the PROC FACTOR procedure with the minimal eigenvalue set at 1 and varimax rotation. The method option applied was principal component analysis, which extracts factors that capture the maximum variation in the data.[36] The components considered were the resting systolic BP and heart rate immediately before the SWT test; the number of shuttles completed; three design variables (0, 1) respectively coding for (a) the early stoppage of exercise on the patient's or supervisor's initiative; (b) moderate to severe breathlessness ($\geq$6 on the 10-point Borg scale), and (c) symptom-limited exercise; and the following ratios (Appendix Fig. S3): systolic BP immediately post- to pre-exercise, systolic BP and heart rate at 5 min post-exercise to the corresponding values immediately postexercise, and the heart rate at 5 min postexercise to pre-exercise (heart rate recovery).

Linear and logistic regression

At the 1-month (baseline) and 9-month (follow-up) visits, ETIF was first related to the urinary peptides derived from non-collagen proteins (n = 198) and in separate runs to peptides from different collagen alpha-chains (n = 152), one peptide at a time, using linear regression models (Fig. 1). Significance levels were corrected by the Benjamini-Hochberg false discovery rate (FDR). The objective of this step was to exclude urinary peptide fragments, which did not retain FDR-corrected significance ($p \leq 0.05$). Covariables were selected a priori based on clinical relevance and known associations with exercise performance, cardiovascular risk and renal function and applied in a hierarchical manner. Models were first adjusted for randomisation group only (not significant in any run[7,8]). For basic adjustment, the covariables added to randomisation group included sex, age, body mass index (BMI), WHR, current smoking (0,1), and history of CHD (0,1), as obtained at screening. Extended adjustment also considered the following screening variables: NT-proBNP (logarithmically transformed), eGFR, the presence of echocardiographic left ventricular hypertrophy (LVH),[37] and the visit-specific use of BP-lowering medicines by drug class: diuretics, β-blockers, inhibitors of the renin-angiotensin system (angiotensin-converting enzyme inhibitors and angiotensin receptor blockers), and vasodilators (calcium-channel blockers and α-blockers). These analyses were conducted on the baseline (discovery) and the 9-month data (replication). The variance inflation factor (VIF) was used to exclude collinearity between the variables in linear models with extended adjustment. For descriptive purposes (Table 1) and in secondary analyses, ETIF was categorised into quartiles (Appendix Table S1 for non-collagen urinary fragments and Appendix Table S2 for collagen-derived peptides). Using logistic regression with three adjustment steps as before, the top quartile (greatest risk of exercise intolerance) was contrasted to the other quartiles.

Elastic net regression

Given the strong correlations between the urinary peptides derived from non-collagen proteins (Appendix Fig. S4) or collagens (Appendix Fig. S5), in the next step of the analyses, ETIF was related to all covariables and to the urinary peptides retained in the foregoing FDR-corrected eliminatory analyses (29 non-collagen peptides [23 proteins] and 31 collagen fragments [16 proteins]) by elastic net regression. The objective here was to construct a parsimonious model predicting exercise intolerance. The λ1 and λ2 regression penalties were determined by 10-fold random cross-validation. This procedure was bootstrapped 1000 times to obtain the final estimates of λ1 and λ2 and the 95% confidence intervals (CI) of the regression coefficients, linking modelled ETIF (ETIFmod) to the covariables and the urinary peptides. As before, these analyses were run separately for peptides derived from non-collagen proteins and collagens. Feature selection in elastic net regression does not rely on multiple independent tests and does not generate p-values. Instead, a multivariable machine learning process, combining penalised elastic net regression and selection based on stability in 1000 bootstrap iterations, was applied.[38] Stability selection

Characteristic	All participants	Quartiles of the exercise performance index				
		Q1	Q2	Q3	Q4	p-value
Quartile limits	—	<−0.6101	<−0.0066	<0.7217	≥0.7217	
Number in group	268	67	67	67	67	
Randomised to spironolactone, n (%)	137 (51.1)	31 (46.3)	32 (47.8)	35 (52.2)	39 (58.2)	0.16
Women, n (%)	61 (22.8)	6 (9.0)	16 (23.9)	18 (26.9)	21 (31.3)	0.0026
Men, n (%)	207 (77.2)	61 (91.0)	51 (76.1)	49 (73.1)	46 (68.7)	
Risk factors						
Current smoking, n (%)	159 (59.3)	39 (58.2)	40 (59.7)	46 (68.7)	34 (50.7)	0.66
Drinking alcohol, n (%)	64 (23.9)	17 (25.4)	24 (35.8)	17 (25.4)	6 (8.9)*	0.012
Office hypertension, n (%)	266 (99.2)	66 (98.5)	67 (100)	66 (98.5)	67 (100)	0.75
Treated hypertension, n (%)	258 (96.3)	63 (94.0)	67 (100)	62 (92.5)	66 (99.5)	0.67
Diuretics, n (%)	32 (11.9)	3 (4.5)	8 (11.9)	9 (13.4)	12 (17.9)	0.023
β-blockers, n (%)	199 (74.2)	51 (76.1)	54 (80.6)	46 (68.7)	48 (71.6)	0.32
RAS inhibitors, n (%)	205 (76.5)	49 (73.1)	54 (80.6)	47 (70.1)	55 (82.1)	0.52
Vasodilators, n (%)	99 (36.9)	20 (29.8)	24 (35.8)	29 (43.3)	26 (38.8)	0.21
Combination therapy, n (%)	194 (72.4)	45 (67.2)	53 (79.1)	46 (68.7)	50 (74.6)	0.67
Diabetes, n (%)	100 (37.3)	17 (25.4)	20 (29.8)	27 (40.3)	36 (53.7)	0.0004
Echocardiographic LVH, n (%)	74 (27.6)	15 (22.4)	15 (22.4)	18 (26.9)	26 (38.8)	0.032
History of CHD, n (%)	219 (81.7)	62 (92.5)	58 (86.6)	54 (80.6)	45 (67.2)	<0.0001
Clinical measurements						
Age (mean), y	73.6 (6.1)	70.2 (4.8)	73.3 (5.9)*	74.2 (5.1)	76.5 (6.6)*	<0.0001
Age (median), y	73.1 (69.0–77.7)	69.9 (66.3–73.7)	72.7 (69.2–78.7)*	74.5 (69.3–77.9)	76.8 (71.2–82.6)*	<0.0001
Body mass index, kg/m^2	28.8 (5.2)	28.2 (4.5)	28.1 (4.1)	29.2 (5.1)	30.0 (6.6)	0.11
Waist-to-hip ratio	0.97 (0.06)	0.96 (0.06)	0.97 (0.06)	0.94 (0.07)	0.99 (0.07)	0.14
Biochemical measurements						
Serum creatinine, µmol/L	90.9 (22.8)	88.9 (17.6)	89.5 (21.6)	90.4 (22.1)	95.0 (28.4)	0.40
eGFR, mL/min/1.73 m^2	66.9 (15.4)	68.0 (13.1)	67.4 (14.5)	67.3 (14.0)	65.0 (19.4)	0.69
Total serum cholesterol, mmol/L	3.91 (1.09)	3.75 (1.03)	3.92 (1.13)	3.84 (0.98)	4.11 (1.20)	0.24
NT-proBNP, pmol/L	24.7 (15.8–39.0)	20.2 (11.6–30.3)	20.8 (11.7–34.1)	26.7 (18.6–41.2)*	33.9 (21.9–48.3)*	<0.0001
Shuttle walk test measurements						
Before exercise						
SBP, mm Hg	143.0 (21.9)	137.8 (16.2)	143.4 (20.4)*	146.3 (21.3)	145.0 (24.3)	0.11
DBP, mm Hg	78.1 (10.2)	78.5 (10.1)	78.4 (10.0)	77.7 (11.8)	77.8 (8.8)	0.95
HR, beats per minute	62.2 (10.3)	58.2 (7.4)	61.9 (9.2)*	61.6 (11.1)	64.5 (9.2)*	<0.0001
During and after exercise						
Number of shuttles (mean), n	49.9 (25.0)	72.9 (18.0)	57.2 (20.8)*	39.9 (17.7)*	27.8 (16.3)*	<0.0001
Number of shuttles (median), n	48 (30–65)	70 (59–87)	57 (44–66)*	38 (28–50)*	25 (15–38)*	<0.0001
SBP increase (median), %	17.4 (6.9–30.1)	33.1 (21.8–48.9)	19.1 (14.1–30.8)*	14.5 (3.7–17.4)*	4.0 (2.8 to −10.9)*	<0.0001
SBP decrease (median), %	15.9 (24.5–8.6)	28.4 (31.9–22.4)	20.0 (25.3–12.6)*	13.7 (17.2–8.0)*	5.7 (12.9 to −0.01)*	<0.0001
HR decrease (median), %	12.6 (23.9–3.8)	23.9 (35.0–13.4)*	16.5 (26.1–6.6)*	12.8 (18.6–3.3)*	3.8 (8.9–0.1)*	<0.0001
HR recovery (median), %	6.6 (0.1–17.6)	25.0 (12.1–36.2)	8.3 (1.6–14.9)*	4.3 (0.1–9.5)	0 (4.6 to −3.3)*	<0.0001
Test stopped, n (%)	249 (92.9)	56 (83.6)	60 (89.6)	66 (98.5)	67 (100)	<0.0001
Symptom-limited test, n (%)	75 (28.0)	4 (6.0)	10 (14.9)	14 (20.9)	47 (70.2)	<0.0001
Breathlessness score (median)	10 (5–15)	13 (9–15)	9 (4–14)*	8 (1–12)	9 (5–19)	0.012

Values are mean (SD), median (interquartile range), or number of participants (%). For NT-proBNP, which is logarithmically transformed, the geometric mean (interquartile range) is given. Hypertension is a blood pressure of ≥140 mm Hg systolic or ≥90 mm Hg diastolic or use of antihypertensive medications. Diabetes is the use of antidiabetic drugs, fasting blood glucose of ≥7.0 mmol/L, random blood glucose ≥11.1 mmol/L, or diabetes documented in practice or hospital records. Use of alcohol is the habitual consumption of alcoholic beverages daily or weekly. Systolic blood pressure (SBP) increase is the percent difference between the resting SBP before the shuttle walk test (SWT) and SBP immediately after the test. SBP/heart rate (HR) decrease is the percent difference between the SBP/HR immediately after the test and 5 min later. HR recovery is the percent difference between the pretest HR and HR 5 min after stopping the test. Test stopped means that either the supervisor or the patient ended the SWT test early, because of a medical indication or symptoms. Symptom-limited test refers to the early termination of the test, because of symptoms perceived as severe by the patient. The breathlessness score is the sum of the scores on the 10-point Borg scale immediately after the test and at 1, 2, 3, and 5 min postexercise. Abbreviations: SBP, systolic blood pressure, DBP, diastolic blood pressure; HR, heart rate, LVH, left ventricular hypertrophy; CHD, coronary heart disease; eGFR, glomerular filtration rate derived from serum creatinine by the Chronic Kidney Disease Epidemiology Collaboration equation; NT-proBNP, N-terminal pro-B-type natriuretic peptide. RAS inhibitors refer to inhibitors of the renin-angiotensin system (angiotensin converting-enzyme inhibitors and angiotensin receptor blockers). Vasodilators include calcium-channel blockers and α-blockers. The p-value denotes the significance of the linear trend across quartiles. An asterisk indicates a significant difference with the left adjacent quartile.

Table 1: **Baseline characteristics of all participants and by quartiles of the exercise performance index.**

intrinsically provides finite-sample control over the expected number of false discoveries.[38] The SHapley Additive exPlanation (SHAP) method,[39] as implemented in R version 4.5.1 (R Core Team, Austria) was applied to rank clinical and UPP features for importance and to visualise their associations with ETIFmod in the whole study population and in individual patients. In a sensitivity analysis, ETIFmod was replaced by the shuttle walk distance covered (SWDmod).

Pathway analysis

Finally, for the urinary peptides consistently related to ETIFmod, a pathway analysis was conducted, using the Gene Ontology (GO) Database (http://www.geneontology.org), Reactome (http://www.reactome.org), and the Kyoto Encyclopaedia of Genes and Genomes (KEGG; http://www.genome.jp/kegg). Enrichment analysis was conducted using the Reactome Pathway Analysis[40] and clusterProfiler[41] packages in R (version 4.5.1). For the GO term analysis, pathways were restricted to the biological processes and molecular function terms. Pathways, not exclusively driven by collagens, were visualised in DotPlots, using R software with q-values corrected for FDR ($p \leq 0.05$).[40]

Role of the funding source

The funding source had no role in study design, data collection, analysis, and interpretation, in the writing of the report, and in the decision to submit the paper for publication.

Results

Patient characteristics

Table 1 lists the descriptive data for the 268 patients, of whom 131 (48.9%) had been randomised to control and 137 (51.1%) to spironolactone. The patients were commonly treated with antihypertensive agents (n = 258; 96.3%) and lipid-lowering drugs (n = 236; 88.1%), mainly statins (n = 231; 86.2%), antiplatelet drugs (n = 203; 75.7%), and antidiabetic agents (n = 93; 34.7%), including insulin in 8 (3.0%) cases. Compared to the 237 randomised HOMAGE patients retained in the master report (Appendix Fig. S1),[27] but not included in the present analyses, the 268 patients analysed had a higher risk profile, as evidenced by the greater use of lipid-lowering drugs, the higher prevalence of antecedent CHD, higher serum cholesterol levels and lower eGFR (Appendix Table S3). Appendix Fig. S6 describes the factor loadings and scoring coefficients of the components of ETIF. Higher ETIF values indicate worse exercise tolerance (Table 1). Across increasing quartiles of ETIF, the proportion of women, the prevalence of diabetes, and the serum level of NT-proBNP increased, whereas proportion of patients with a history of CHD decreased (Table 1). ETIF was positively correlated with age and inversely with the walking distance covered (Appendix Fig. S7). The median left ventricular ejection fraction was 62.4% (IQR, 57.1–65.7%; 5th–95th percentile interval, 50.2–71.8%) at baseline (n = 224) and 62.7% (IQR, 57.0–66.4%; 5th–95th percentile interval, 51.9–71.9%) at the 9-month visit (n = 180). Median eGFR was 62.1 mL/min/1.73 m^2 (IQR, 57.0–78.3 mL/min/1.73 m^2; 5th–95th percentile interval, 44.2–91.0 mL/min/1.73 m^2).

Association of ETIF with urinary peptides

In analyses of the baseline and 9-month data with FDR correction of significance and adjustment for randomisation group, ETIF was associated with 29 of 198 non-collagen urinary peptides (Table 2) and with 31 of 152 collagen peptides (Table 3). Of these 29 non-collagen peptides and 31 collagen fragments, the median proportion of imputed values was 37.3% (IQR, 17.7–58.9%) at baseline and 38.1% (IQR, 15.7–62.1%) at follow-up. Positive partial regression coefficients indicated worse exercise tolerance and negative coefficients better exercise tolerance. The regression coefficients at baseline and month 9 showed complete directional consistency. With increasing adjustment, associations of ETIF with the UPP weakened. The parent non-collagen proteins of eight urinary peptides that kept FDR-corrected significance in relation to ETIF at baseline, follow-up or both were ADAM18, AHSG, ANXA1, FXYD2, KCNK3, MUC12, SECTM1, and SNX9 (Table 2); the seven corresponding parent collagen alpha chains were COL1A2, COL3A1, COL4A1, COL5A2, COL6A1, COL11A2, and COL15A1 (Table 3). To search for collinearity, three collagen fragments were selected, because they were significantly associated with ETIF at baseline and follow-up and at each step of the increasing adjustment. These collagen fragments (Table 3) were derived from COL3A1 (laboratory identification number, 18,864), COL1A2 (18,867), and COL6A1 (17,447). VIF associated with all variables in the six models ranged from 1.04 to 1.66; VIF associated with treatment allocation from 1.08 to 1.09; and VIF associated with the urinary peptides from 1.08 to 1.18.

In the categorical analyses of ETIF in relation to non-collagen (Appendix Table S1) and collagen-derived (Appendix Table S2) urinary peptides, contrasting the top *vs.* the other quartiles, standardised odds ratios greater or smaller than unity consistently matched the positive and negative regression coefficients in the continuous analyses (Tables 2 and 3). The 29 non-collagen and the 31 collagen-derived urinary peptides, respectively representing 23 and 16 parent proteins, along with clinical and biochemical covariables were carried through to the next assessment step (Fig. 1), the

Peptides		Baseline			Month 9		
Symbol	ID	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
ADAM18	12,766	0.230† (0.099, 0.362)	0.123* (0.006, 0.241)	0.128* (0.009, 0.246)	0.216† (0.087, 0.345)	—	—
AHSG	05802	0.214* (0.070, 0.359)	0.122* (0.011, 0.303)	0.139* (0.008, 0.269)	0.259† (0.102, 0.416)	0.149* (0.006, 0.291)	0.179* (0.035, 0.322)
AHSG	07656	0.148* (0.012, 0.285)	—	—	0.176* (0.037, 0.316)	—	0.138* (0.011, 0.265)
ANXA1	04435	0.156* (0.032, 0.281)	—	0.123* (0.011, 0.235)	0.253† (0.125, 0.380)	0.172† (0.057, 0.288)	0.176† (0.062, 0.290)
ATG12	04609	−0.172* (−0.305, −0.038)	—	—	−0.257† (−0.389, −0.126)	−0.189† (−0.306, −0.071)	−0.150* (−0.270, −0.029)
FXYD2	00939	−0.128* (−0.249, −0.007)	—	—	—	—	—
FXYD2	02740	—	—	—	−0.198* (−0.308, −0.027)	−0.147* (−0.269, −0.025)	−0.137* (−0.261, −0.014)
FXYD2	05387	−0.181* (−0.301, −0.061)	−0.173† (−0.279, −0.068)	−0.167† (−0.276, −0.059)	−0.241† (−0.360, −0.122)	−0.173† (−0.279, −0.066)	−0.149† (−0.258, −0.041)
GAPDH	15,365	0.195* (0.059, 0.332)	—	—	0.162* (0.029, 0.294)	—	—
HBB	05269	−0.157* (−0.301, −0.013)	—	—	—	—	—
KCNB1	10,519	—	—	—	0.134* (0.008, 0.260)	—	—
KCNK3	13,874	−0.211* (−0.355, −0.068)	—	−0.146* (−0.271, −0.021)	−0.195* (−0.338, −0.053)	−0.123* (−0.249, −0.002)	−0.155* (−0.239, −0.001)
MB	04940	0.171* (0.038, 0.304)	—	—	—	—	—
MB	11,919	0.156* (0.016, 0.297)	—	—	—	—	—
MGP	00340	0.163* (0.042, 0.284)	—	—	0.131† (0.010, 0.252)	—	—
MGP	01132	0.218† (0.098, 0.339)	0.121* (0.012, 0.230)	—	0.186* (0.070, 0.302)	—	—
MUC12	17,804	−0.306‡ (−0.429, −0.183)	−0.181† (−0.297, −0.066)	−0.167† (−0.282, −0.051)	−0.231† (−0.356, −0.107)	−0.163† (−0.274, −0.051)	−0.159† (−0.270, −0.047)
MYOCD	12382	−0.140* (−0.265, −0.014)	−0.109* (−0.220, −0.001)	—	—	—	—
RAF1	12740	0.181* (0.041, 0.321)	—	—	0.219† (0.075, 0.363)	—	—
SECTM1	10594	0.189* (0.047, 0.331)	0.133* (0.040, 0.989)	0.149* (0.024, 0.275)	0.206† (0.065, 0.348)	0.155* (0.030, 0.280)	0.164* (0.039, 0.290)
SETD1B	03624	−0.254† (−0.397, −0.111)	−0.157* (−0.283, −0.030)	−0.148* (−0.273, −0.022)	−0.165* (−0.310, −0.020)	—	—
SNX9	13327	−0.136* (−0.262, −0.010)	−0.116* (−0.222, −0.004)	—	−0.158* (−0.286, −0.031)	−0.139* (−0.250, −0.028)	−0.144* (−0.253, −0.034)
SOD1	06033	—	—	—	−0.175* (−0.303, −0.048)	—	—
SSP1	15313	0.265† (0.127, 0.404)	—	—	0.202† (0.064, 0.341)	—	—
TACC3	13243	0.159* (0.017, 0.301)	—	—	0.157* (0.013, 0.302)	—	—
TMSB4X	20740	0.216† (0.086, 0.346)	—	—	0.240† (0.110, 0.371)	—	—
TTN	12157	—	—	—	−0.152* (−0.299, −0.005)	—	—
UMOD	08455	−0.134* (−0.267, −0.003)	—	—	—	—	—
UMOD	10189	—	—	—	−0.136* (−0.259, −0.014)	—	—

The analysis includes 268 patients, who performed the shuttle walk test at screening and had urinary proteomic data at baseline (discovery) and at the 9-month follow-up visit (replication). Exercise performance refers to the standardised factor (ETIF) summarising measurements taken immediately before, during, and after the shuttle walk test. Higher values of this factor indicate worse exercise tolerance. Values are standardised regression coefficients given with 95% confidence interval. Urinary peptides are rank-normalised and are listed with their amino-acid sequence in Appendix Table S4. Model 1 is adjusted for randomisation group only. Model 2 additionally accounts for sex, age, body mass index, waist-to-hip ratio, current smoking, and history of coronary heart disease. Model 3 furthermore includes N-terminal pro-B-type natriuretic peptide (logarithmically transformed), the glomerular filtration rate derived from serum creatinine by the Chronic Kidney Disease Epidemiology Collaboration equation, the presence of echocardiographic left ventricular hypertrophy, and current treatment with diuretics, β-blockers, inhibitors of the renin-angiotensin system (angiotensin-converting enzyme inhibitors or angiotensin receptor blockers), or vasodilators (calcium-channel blockers or α-blockers). Significance of the partial regression coefficients with adjustment for the false discovery rate: *$p \leq 0.05$; †$p \leq 0.01$; ‡$p \leq 0.001$. An emdash indicates that there is no significant association between exercise performance (ETIF) and the urinary peptide.

Table 2: **Association of exercise performance with urinary peptides derived from proteins other than collagens.**

Peptides			Baseline			Month 9		
Symbol	ID	Mass (Da)	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
COL1A1	19111	3802.7	0.198† (0.064, 0.332)	—	—	0.187* (0.055, 0.319)	—	—
COL1A1	19152	3818.7	0.162* (0.016, 0.308)	—	—	0.219* (0.073, 0.366)	—	—
COL1A1	19279	3871.8	−0.214† (−0.333, −0.095)	—	—	−0.195† (−0.312, −0.079)	—	—
COL1A1	20010	4252.0	−0.193† (−0.315, −0.072)	—	—	−0.129* (−0.250, −0.008)	—	—
COL1A2	18867	3719.7	−0.215† (−0.337, −0.093)	−0.150* (−0.260, −0.040)	−0.147* (−0.257, −0.037)	−0.181* (−0.305, −0.056)	−0.122* (−0.232, −0.012)	−0.112* (−0.222, −0.002)
COL1A2	18907	3735.7	−0.196† (−0.318, −0.074)	−0.112* (−0.223, −0.002)	—	−0.157* (−0.283, −0.031)	—	—
COL3A1	18864	3718.7	−0.223† (−0.344, −0.103)	−0.132* (−0.242, −0.021)	−0.134* (−0.247, −0.022)	−0.229† (−0.349, −0.109)	−0.127* (−0.238, −0.016)	−0.122* (−0.232, −0.012)
COL3A1	20065	4289.9	−0.129* (−0.252, −0.007)	—	—	−0.138* (−0.258, −0.018)	—	—
COL3A1	20109	4322.0	0.150* (0.012, 0.289)	—	—	0.238† (0.098, 0.378)	—	—
COL4A1	08436	1821.8	−0.139* (−0.275, −0.003)	−0.133* (−0.251, −0.015)	−0.122* (−0.240, −0.004)	−0.139* (−0.164, −0.015)	−0.140* (−0.271, −0.010)	−0.135* (−0.265, −0.006)
COL4A1	14943	2777.3	0.126* (0.001, 0.252)	—	—	0.264† (0.136, 0.392)	0.166* (0.049, 0.284)	0.175* (0.059, 0.291)
COL4A3	15360	2847.3	−0.137* (−0.259, −0.014)	—	—	−0.156* (−0.279, −0.033)	—	—
COL4A5	16342	3048.4	−0.212† (−0.347, −0.077)	−0.136* (−0.260, −0.012)	—	—	—	—
COL5A1	18060	3463.6	−0.148* (−0.273, −0.022)	—	—	—	—	—
COL5A2	15320	2838.4	0.178† (0.056, 0.300)	—	—	0.128* (0.008, 0.248)	—	—
COL5A2	20387	4523.1	−0.303† (−0.448, −0.158)	−0.221† (−0.349, −0.093)	−0.207* (−0.336, −0.079)	−0.192* (−0.346, −0.037)	−0.158* (−0.291, −0.025)	−0.174* (−0.306, −0.042)
COL6A1	17131	3231.4	−0.144* (−0.267, −0.022)	—	—	−0.146* (−0.270, −0.022)	—	—
COL6A1	17447	3304.4	0.256† (0.134, 0.377)	0.138* (0.027, 0.248)	0.123* (0.012, 0.235)	0.258† (0.140, 0.376)	0.132* (0.022, 0.242)	0.122* (0.013, 0.232)
COL6A1	20326	4468.0	−0.201† (−0.321, −0.081)	—	—	−0.196† (−0.318, −0.074)	—	—
COL7A1	20262	4414.0	−0.253† (−0.377, −0.129)	−0.158* (−0.273, −0.042)	−0.133* (−0.249, −0.016)	−0.180* (−0.306, −0.054)	−0.140* (−0.252, −0.028)	—
COL11A1	17283	3265.4	−0.131* (−0.252, −0.009)	−0.107* (−0.213, −0.001)	—	−0.133* (−0.256, −0.009)	−0.116* (−0.224, −0.008)	—
COL11A2	11726	2259.0	0.158* (0.036, 0.280)	—	—	0.163* (0.050, 0.279)	—	—
COL11A2	12367	2339.0	−0.240† (−0.387, −0.094)	−0.151* (−0.283, −0.021)	−0.141* (−0.271, −0.011)	−0.162* (−0.314, −0.011)	—	—
COL11A2	17650	3350.5	−0.150* (−0.273, −0.026)	−0.155* (−0.225, −0.004)	—	−0.131* (−0.261, −0.002)	−0.126* (−0.238, −0.014)	−0.119* (−0.231, −0.008)
COL11A2	17871	3411.6	−0.155* (−0.293, −0.016)	—	—	−0.143* (−0.280, −0.006)	−0.121* (−0.240, −0.003)	—
COL11A2	19210	3839.8	0.241† (0.121, 0.361)	0.174* (0.067, 0.280)	0.167* (0.056, 0.277)	0.204† (0.079, 0.330)	—	—
COL15A1	02352	1161.5	−0.161* (−0.286, −0.037)	−0.130* (−0.239, −0.021)	−0.118* (−0.228, −0.008)	−0.157* (−0.220, −0.032)	−0.150* (−0.261, −0.039)	−0.130* (−0.241, −0.019)
COL18A1	05725	1529.7	−0.150* (−0.291, −0.008)	−0.125* (−0.248, −0.002)	—	—	—	—
COL18A1	06356	1600.7	—	—	—	0.154* (0.030, 0.279)	—	—
COL26A1	00600	935.5	—	—	—	0.143* (0.004, 0.283)	—	—
COL28A1	07469	1716.8	−0.155* (−0.276, −0.033)	−0.125* (−0.234, −0.017)	−0.118* (−0.226, −0.010)	−0.131* (−0.256, −0.006)	—	—

The analysis includes 268 patients, who performed the shuttle walk test at screening and had urinary proteomic data at baseline (discovery) and at the 9-month follow-up visit (replication). Exercise performance refers to the standardised factor (ETIF) summarising measurements taken immediately before, during, and after the shuttle walk test. Higher values of this factor indicate worse exercise tolerance. Values are standardised regression coefficients given with 95% confidence interval. Urinary peptides are rank-normalised and are listed with their amino-acid sequence in Appendix Table S4. Model 1 is adjusted for randomisation group only. Model 2 additionally accounts for sex, age, body mass index, waist-to-hip ratio, current smoking, and history of coronary heart disease. Model 3 furthermore includes N-terminal pro-B-type natriuretic peptide (logarithmically transformed), the glomerular filtration rate derived from serum creatinine by the Chronic Kidney Disease Epidemiology Collaboration equation, the presence of echocardiographic left ventricular hypertrophy, and current treatment with diuretics, β-blockers, inhibitors of the renin-angiotensin system (angiotensin-converting enzyme inhibitors or angiotensin receptor blockers), or vasodilators (calcium-channel blockers or α-blockers). Significance of the partial regression coefficients with adjustment for the false discovery rate: *$p \leq 0.05$; †$p \leq 0.01$; ‡$p \leq 0.001$. An emdash indicates that there is no significant association between exercise performance (ETIF) and the urinary peptide.

Table 3: **Association of exercise performance with urinary peptides derived from collagens.**

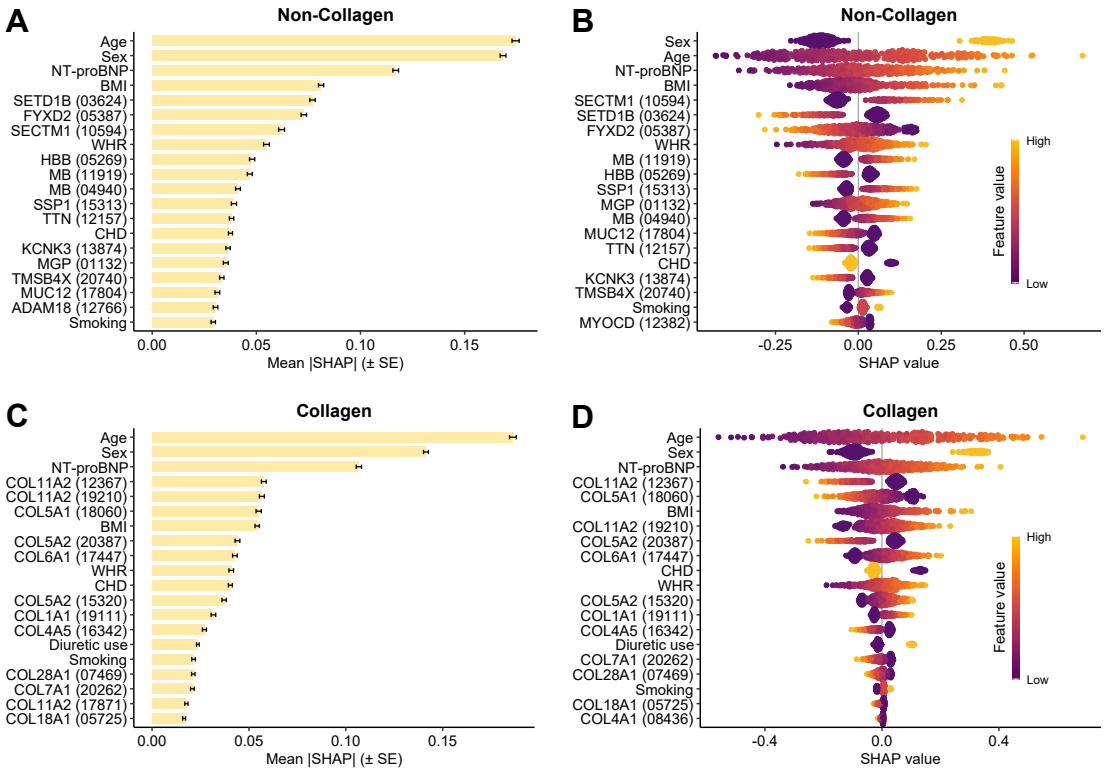

Fig. 2: Importance of features in their association with ETIFmod in the whole study population and in individual patients in analyses involving non-collagen (A, B) and collagen-derived (C, D) urinary peptides. Panel (A) and panel (C) rank the features for non-collagen and collagens according to their importance and panel (B) and panel (D) show the individual data points for each feature. SHAP values indicate how much each feature contributes to ETIFmod, facilitating the understanding of the elastic net regression model. The non-collagen peptides and their parent proteins are listed in Appendix Table S4 and the collagen-derived peptides and their parent proteins in Appendix Table S5. Abbreviations: NT-proBNP, N-terminal pro-B-type natriuretic peptide (logarithmically transformed), BMI, body mass index; WHR, waist-to-hip ratio; CHD, history of coronary heart disease.

elastic net regression to generate ETIF modelled (ETIFmod).

Elastic net regression

The feature stability of ETIFmod across 1000 bootstrapped runs considering both the baseline and the 9-month UPP data are summarised in Appendix Table S6 for non-collagen peptides and in Appendix Table S7 for collagen fragments. The clinical and biochemical variables showing great consistency in the bootstrapped runs both at baseline and at 9 months included sex, age, BMI, WHR, serum NT-proBNP at screening (logarithmically transformed), history of CHD, current smoking, and the use of diuretics, β-blockers, and inhibitors of the renin-angiotensin system (Appendix Table S6 and S7). Fig. 2 shows the relative importance of the ETIFmod features in the whole study population and in individual patients in analyses involving non-collagen and collagen-derived urinary peptides. Panel (a) and panel (c) rank the features for non-collagens and collagens according to their importance and panel (b) and panel (d) show the individual data points for each feature. Given the features listed in Fig. 2, there was large interindividual variability in the association of ETIFmod with the clinical and biochemical variables and the UPP components (Appendix Fig. S8). Urinary peptides selected in at least 50% of the bootstrapped runs indicate feature stability. The corresponding non-collagen parent proteins included FXYD2, GAPDH, HBB, KCNB1, MB, MYOCD, SECTM1, SNX9, TACC3, TMSB4X, and TTN (Fig. 3); the stable parent collagens were COL5A1, COL6A1, COL11A2, and COL28A1 (Fig. 3).

Elastic net regression was also applied to relate one component of ETIFmod, i.e., the walking distance covered (SWDmod) to the same variables as ETIFmod. The proteins replicated in the SWDmod analysis included FXYD2, MB, MYOCD, and SNX9 for the non-collagens (Appendix Table S8) and COL11A2 and COL28A1 for the collagens (Appendix Table S9). The proteins contributing to ETIFmod, SWDmod, or both are graphically summarised in Appendix Fig. S8.

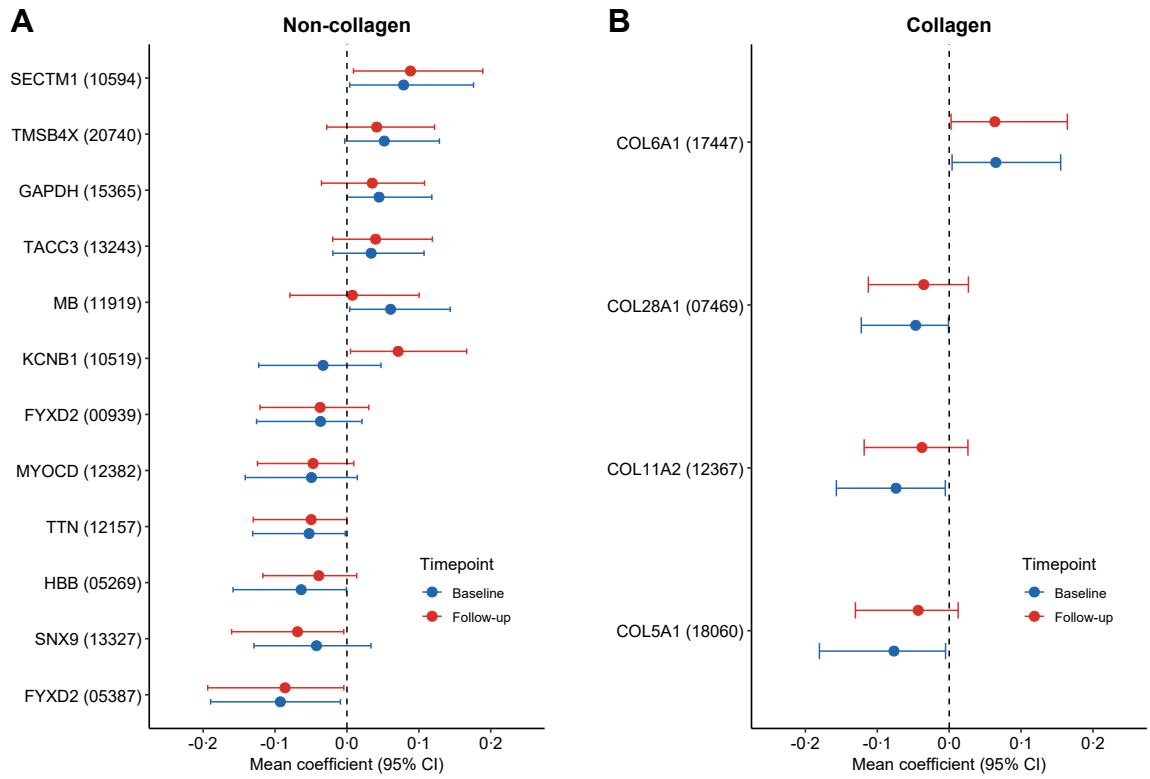

Fig. 3: Parent proteins of the non-collagen (A) and collagen-derived peptides (B), which show feature stability in 1000 times bootstrapped runs of the elastic net regression at baseline (discovery) and at the 9-month follow-up (replication). Features selected in ≥50% of runs are stable. Bracketed numbers following the protein symbol are the laboratory identification numbers of sequenced peptides identifying the parent proteins, as listed in Appendix Table S4 for non-collagens and in Appendix Table S5 for collagens.

Pathway analysis

Pathway analysis identified 25 (GO biological process and molecular function terms), 24 (Reactome), and two (KEGG) pathways (FDR-corrected p-value ≤0.05). Pathways driven by non-collagen proteins and collagens (Fig. 4 and Appendix Table S10) were involved in oxygen homoeostasis (transport and exchange pathways), oxidative stress regulation (antioxidant activity, peroxidase functions, and response to reactive oxygen species), and metabolic and developmental processes (including ATP and nucleotide metabolism, platelet activity, cytoskeletal organisation of muscle cells). Of the overrepresented pathways, 28 were exclusively driven by collagens (Appendix Table S11) and were mainly related to collagen biology, dynamics of the extracellular matrix (ECM), and cell signalling.

Discussion

The key finding of this article was that in patients at high HF risk exercise intolerance as captured by ETIFmod was associated with 11 non-collagen proteins, FXYD2, GAPDH, HBB, KCNB1, MB, MYOCD, SECTM1, SNX9, TACC3, TMSB4X, and TTN, and with four collagen-derived alpha-chains, COL5A1, COL6A1, COL11A2, and COL28A1. Relevant pathways involving these proteins included oxygen homoeostasis, oxidative stress regulation, metabolic and developmental processes, and the cytoskeletal organisation of muscle cells. A sensitivity analysis with SWDmod as outcome, one of 10 components summarised in ETIFmod, was still associated with FXYD2, MB, MYOCD, SNX9, COL11A2, and COL28A1.

The application of proteomics in experimental studies[14–16] substantially deepened insight in the pathophysiology of exercise tolerance. Submaximal exercise capacity and its trainability are for about 40% heritable.[42] A seminal study[15] assessed submaximal exercise capacity before and after training in relation to 2.5×10^6 single nucleotide polymorphisms, transcriptomics, plasma proteomics, and metabolomics in 446 white participants from the HERITAGE family study. The leading omics features were retested, involving 117 genes related to submaximal exercise capacity and 90 related to trainability in genetically engineered mice. The most prominent genes and pathways contributing to human variability in submaximal exercise capacity as compared with trainability differed markedly, but the

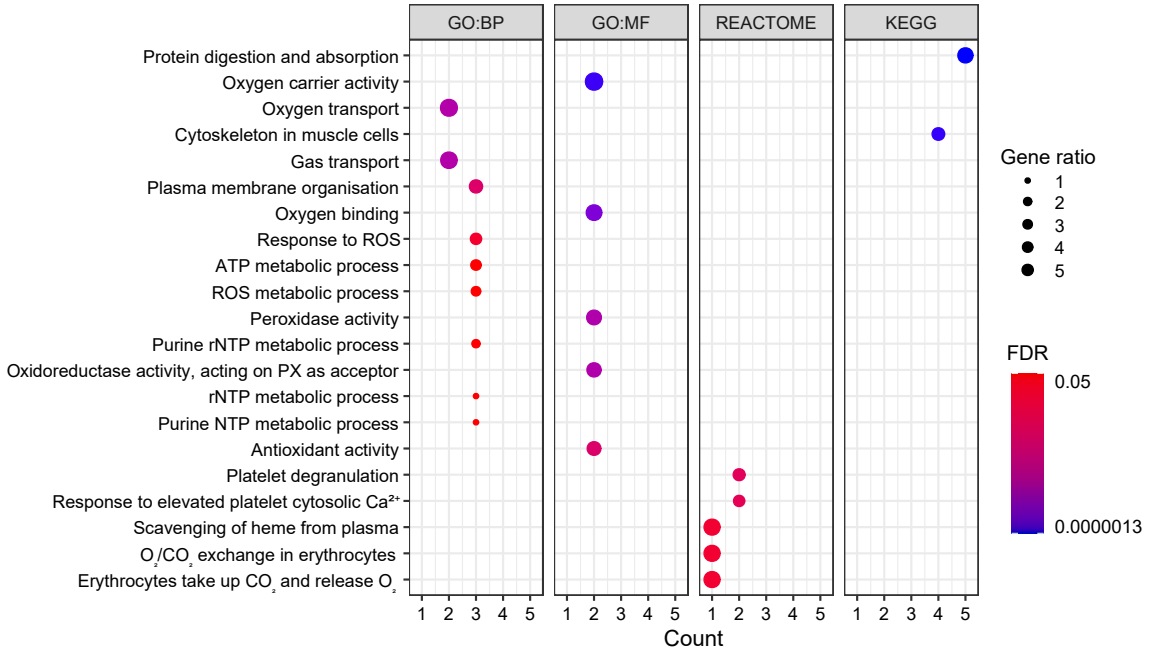

Fig. 4: Pathway analysis using the Gene Ontology, Reactome, and Kyoto Encyclopaedia of Genes and Genomes Databases. The presented pathways are driven by non-collagen proteins and collagens. FDR refers to p-values accounting for the false discovery rate. Abbreviations: GO:BP, gene ontology biological process; GO:MF, gene ontology molecular function; ROS, reactive oxygen species; ATP, adenosine triphosphate; rNTP, ribonucleoside triphosphate; NTP, nucleoside triphosphate; PX, peroxidase; RBC, red blood cell.

key genes and pathways were involved in the morphology and metabolism of skeletal muscles and oxygen transport.[15] These findings are in keeping with the current pathway analysis pointing to the importance of oxygen homoeostasis and the muscular cytoskeletal organisation. A recent HERITAGE report included 654 healthy individuals from families of European or African descent, who completed a 20-week non-randomised exercise programme and underwent CPET and plasma proteomics.[18] Training modulated hundreds of circulating proteins, 40% of which were known to be secreted. The investigators discovered novel proteins, which play a role in angiogenesis, iron homoeostasis, and the turnover of the ECM.[18]

The evidence from the studies referenced above is of great value in understanding the interactive and networked role of multiple genetic, transcriptomic, proteomic, and metabolic biomarkers in the physiology of exercise endurance and trainability. However, to some extent, the complexity of these studies compounds their clinical interpretability.[14–16] This observation justified the current strategy to limit the analyses of the non-collagen and collagen-derived peptidome to parent proteins with known function and expression in relevant tissues. Among the selected non-collagen proteins (Fig. 3), HBB, MB, MYOCD, FXYD2, TMSB4X, TTN, and KCNB1 clearly meet these criteria. HBB and MB modulate the oxygen supply to the mitochondria in skeletal muscle. Endurance training facilitates myoglobin (MB) oxygen desaturation.[43] MYOCD is the most important coactivator of serum response factor (SRF). This protein plays a critical role in the development of cardiac myocytes and vascular smooth muscle cells.[44] Binding of MYOCD to SRF transcriptionally activates a variety of downstream muscle-specific genes, such as *Sm22a, Acta2, Myh11.* MYOCD expression results in a contractile and differentiated smooth muscle phenotype, whereas deficient MYOCD expression leads to a dedifferentiated phenotype, a hallmark in atherosclerosis.[44] In humans, there are seven tissue-specific FXYD proteins, of which FXYD2 (the γ-subunit) is expressed as two splice variants, the γa and γb subunits in the kidney.[45] Both γ-subunit variants are primarily expressed in the thick ascending limb. FXYD4 is exclusively expressed in the distal nephron and enhances the affinity of Na^+/K^+-ATPase for sodium. FXYD2 and FXYD4 fine-tune sodium affinity along the nephron and contribute via sodium reabsorption to the regulation of the circulating intravascular volume.[45] TMSB4X plays a role in actin polymerisation. This gene is located on the X-chromosome, but escapes X inactivation and has a homolog on chromosome Y. TMSB4X is expressed equivalently in the lung tissue of women and men, but has male-biased expression in the skin, kidney and adipose tissue.[46] Actin filaments are a major component of the cytoskeleton, responsible for

cell shape, movement, and intracellular transport, and in association with myosin represent the machinery of muscle contraction. TTN is a key component in the assembly and functioning of striated muscles.[47] By providing connections between individual microfilaments, TTN contributes to the balance of forces generated by the sarcomere. KCNB1 is a voltage-gated potassium channel expressed in excitable cells and is a major contributor to the slowly inactivating delayed-rectifier voltage-gated current in neurons, cardiomyocytes, smooth muscle cells, and pancreatic β-cells. KCNB1 regulates the duration of the action potential of cardiomyocytes and ventricular repolarisation (QT interval).[48] In pancreatic β-cells, KCNB1 limits the inward calcium current and insulin secretion in response to glucose. GAPDH has glyceraldehyde-3-phosphate dehydrogenase and nitrosylase activities, thereby playing a role in glycolysis and nuclear functions, respectively. Glyceraldehyde-3-phosphate dehydrogenase is a key enzyme in glycolysis that catalyses the first step of the pathway by converting D-glyceraldehyde 3-phosphate (G3P) into 3-phospho-D-glyceroyl phosphate.[49] The nuclear functions, including transcription, RNA transport, DNA transport and apoptosis,[49] are probably due to the nitrosylase activity that mediates cysteine S-nitrosylation of nuclear target proteins. Collagen is the major structural protein in the skeletal muscular ECM; it accounts for 1–10% of muscle mass dry weight.[17] Collagen types I, III, IV, V, VI, XI, XII, XIV, XV, and XVIII collagen are expressed during skeletal muscle development, but fibrillar types I and III predominate in adult muscle.[17]

The current study must be interpreted within the context of its limitations. First, HOMAGE recruited patients at high HF risk, in which insulin resistance and adiposity are major drivers of exercise intolerance. The proteins to be targeted to mitigate these metabolic effect may be different from those identified by UPP. Nonetheless, KCNB1 was among the selected proteins and in pancreatic β-cells, inhibits insulin secretion in response to glucose. Serum insulin was not measured in HOMAGE, and insulin resistance could not be accounted for in the multivariable-adjusted net elastic regression. Second, treatment allocation did not affect exercise tolerance.[7,8] However, if spironolactone meaningfully shifts the urinary peptidome,[29] collinearity between treatment and urinary peptides might bias the interpretation, but analysis of VIF for three exemplary collagen fragments did not produce any evidence that the prerequisite in regression analysis, i.e., independence of the explanatory variables, was violated. Third, given the relatively small sample size and the short follow-up, translating the SWT results and, by extension, the clinical significance of ETIFmod in hard adverse outcomes is impossible. However, there is a vast literature showing that exercise intolerance is a predictor of mortality[50,51] and cardiovascular[51] and cardiac[52] endpoints. Fourth, women were underrepresented in this report, limiting its generalisability. Fifth, trainability was not among the objectives of the current study. Sixth, as evidenced by the distribution of left ventricular ejection fraction at baseline and follow-up, patients with acute or chronic HF were excluded from the trial. This makes it difficult to conceptualise the current findings in terms of HF with reduced or preserved ejection fraction. However, given the distribution of age (mean, 73.6 years) and sex (77.2% men) and the presence of underlying CHD, HF with reduced ejection fraction might well be the most likely long-term outcome in patients progressing to overt HF.[53] Seventh, the current observations identified only one protein, FXYD2, which is expressed in the renal tubules and enhances the affinity of Na^+/K^+-ATPase for sodium and probably affects exercise tolerance via regulation of the circulating volume. FXYD2 is not known as primary driver of renal disease. On the other hand, the cardiorenal syndrome encompasses a spectrum of disorders involving both heart and kidneys, in which acute or chronic dysfunction in one organ induces acute or chronic dysfunction in the other with excessive fibrosis as hallmark of organ damage.[54] Additionally, chronic kidney disease impairs the quality and quantity of striated muscles.[55] The current analysis did not reveal any protein that might be involved in these interactions between heart and kidney, probably because the patients did not have chronic kidney disease, as highlighted by the distribution of eGFR. Finally, an extensive literature search did not identify any suitable external validation dataset, because the current UPP analysis of exercise intolerance is based on the urinary peptidome determined by CE-MS. In one small study (n = 10),[26] urine samples were collected before and after simulated firefighting tasks and analysed by liquid chromatographic analysis and proteins were identified via a proprietary spectra library. After the test, 360 proteins were upregulated and 265 down-regulated. The design of this small study[26] with an intervention running over less than 5 min cannot be compared with the HOMAGE design. Indeed, in the current study, the baseline and 9-month data were analysed completely independently from one another and showed consistency. Admittedly, the baseline and follow-up datasets consisted of the same individuals, so that what has been shown might be regarded as internal temporal consistency over 9 months rather than as a replication in the strict sense of the word.

In conclusion, this article demonstrated that SWT, when combined with UPP, is informative for disentangling the mechanisms underlying exercise intolerance. In a proof-of-concept analysis, a composite index of exercise intolerance (ETIF) was related to 29 peptides from 23 non-collagen proteins and 31 peptides from 16 collagens. Via elastic net regression, using ETIFmod as dependent variable, 11 non-collagen parent proteins

and four collagens identified by sequencing of the urinary peptides were associated with exercise intolerance via relevant pathways involving oxygen homoeostasis, regulation of oxidative stress, metabolic and developmental processes, and the cytoskeletal organisation in muscle.

Contributors

FZ coordinated the European Union Seventh Framework Program HOMAGE grant. FZ, JGFC, and JAS are lead investigators who conceived the HOMAGE trial research idea and methodology and had access to all data. FZ and JAS acquired funding. NG constructed the HOMAGE master database. JS, AL, and HM supervised the UPP analyses. FFW, YLY, DYZ and JAS constructed the database for the current analysis. DWA and JAS did the statistical analysis. DSM and TSN generated the pathway analysis. CHL and JAS searched the literature. CHL and DWA produced the graphics. CHL, DWA, and JAS wrote the first draft of the manuscript. PP, JAJV, FZA, PR, JP, SH, JJC, NG, and ALC coordinated patient recruitment and follow-up. DSM, PV, YL, and TSN contributed expertise in interpreting molecular pathways. CHL and JAS accessed and verified the underlying data. All co-authors critically revised the successive drafts of the manuscript, approved the final version, and collectively decided to submit the manuscript for publication.

Data sharing statement

Data sharing is ethically not allowed, because it was not covered by the informed written consent form signed by participants and because of the European Data Protection Regulation (reference 28).

Declaration of interests

Justyna Siwy and Agnieszka Latosinska are employees of Mosaiques-Diagnostics GmbH, Hannover, Germany. Harald Mischak is the co-founder and co-owner of Mosaiques- Diagnostics GmbH. The other authors declare no conflict of interest.

Acknowledgements

The authors are indebted to the many investigators, who were involved in HOMAGE. Their names are listed in the Appendix page 2. This article was submitted for publication on their behalf. HOMAGE was funded by the European Union Seventh Framework Program. Omron Healthcare, Kyoto, Japan provided a non-binding grant to the Non-Profit Research Association Alliance for the Promotion of Preventive Medicine (APPREMED), Mechelen, Belgium. The Guangci Laureate Professorship of Jan A. Staessen is supported by the Guangci Deep Mind Project of Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China.

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.ebiom.2026.106406.

References

1 von Haehling S, Arzt M, Doehner W, et al. Improving exercise capacity and quality of life using non-invasive heart failure treatments: evidence from clinical trials. *Eur J Heart Fail*. 2021;23:92–113.
2 Martens P, Augusto SN, Finet JE, Tang WHW. Distinct impact of noncardiac comorbidities on exercise capacity and functional status in chronic heart failure. *JACC Heart Fail*. 2023;11:1365–1376.
3 Singh R, Aggarwal D, Dutta K, Jaggi S, Sodhi MK, Saini V. Assessment of the feasibility of 1-min sit-to-stand test in evaluating functional exercise capacity in interstitial lung disease patients. *J Exerc Rehabil*. 2023;19:363–369.
4 Balady GJ, Arena R, Sietsema K, et al. Clinician's guide to cardiopulmonary exercise testing in adults: a scientific statement from the American Heart Association. *Circulation*. 2010;122:191–225.
5 Pulz C, Diniz RV, Alves ANF, et al. Incremental shuttle and six-minute walking tests in the assessment of the functional capacity in chronic heart failure. *Can J Cardiol*. 2008;24:131–135.
6 Parreira VF, Janaudis-Ferreira T, Evans RA, Mathur S, Goldstein RS, Brooks D. Measurement properties of the incremental shuttle walk test: a systematic review. *Chest*. 2014;145:1357–1369.
7 Wei FF, Mariottoni B, An DW, et al. Effects of spironolactone on exercise blood pressure in patients at increased risk of developing heart failure: report from the HOMAGE trial. *Hypertens Res*. 2024;47:3225–3236.
8 Wei FF, Mariottoni B, An DW, et al. Heart rate reactivity, recovery, and endurance of the incremental shuttle walk test in patients prone to heart failure. *ESC Heart Fail*. 2024;11:4116–4126.
9 Cohen JB, Schrauben SJ, Zhao L, et al. Clinical phenogroups in heart failure with preserved ejection fraction: detailed phenotypes, prognosis, and response to spironolactone. *J Am Coll Cardiol Heart Fail*. 2020;8:172–184.
10 Latosinska A, Siwy J, Mischak H, Frantzi M. Peptidomics and proteomics based on CE-MS as a robust tool in clinical application: the past, the present, and the future. *Electrophoresis*. 2019;40:2294–2308.
11 Mavrogeorgis E, Mischak H, Latosinska A, Siwy J, Jankowski V, Jankowski J. Reproducibility evaluation of urinary peptide detection using CE-MS. *Molecules*. 2021;26:7260.
12 Hesketh SJ, Stansfield BN, Stead CA, Burniston JG. The application of proteomics in muscle exercise physiology. *Exp Rev Proteomics*. 2020;17:813–825.
13 Murach KA, Bagley JR. A primer on global molecular responses to exercise in skeletal muscle: omics in focus. *J Sport Health Sci*. 2025;15:101029.
14 De Nardo W, Miotto PM, Bayliss J, et al. Proteomic analysis reveals exercise training induced remodelling of hepatokine secretion and uncovers syndecan-4 as a regulator of hepatic lipid metabolism. *Mol Metabol*. 2022;60:101491.
15 Hota M, Barber JL, Ruiz-Ramie JJ, et al. Omics-driven investigation of the biology underlying intrinsic submaximal capacity and its trainability. *Physiol Genomics*. 2023;55:517–543.
16 Gaitán JM, Moon HY, Stremlau M, et al. Effects of aerobic exercise training on systemic biomarkers and cognition in late middle-aged adults at risk for Alzheimer's disease. *Front Endocrinol*. 2021;12:660181.
17 Gillies AR, Lieber RL. Structure and function of the skeletal extracellular matrix. *Muscle Nerve*. 2011;44:318–331.
18 Robbins JM, Rao P, Deng S, et al. Plasma proteomic changes in response to exercise training are associated with cardiorespiratory fitness adaptations. *JCI Insight*. 2023;8:e165867.
19 Nieman DC, Groen AJ, Pugachev A, Vacco G. Detection of functional overreaching in endurance athletes using proteomics. *Proteomes*. 2018;6:33.
20 Dorian D, Gustafson D, Quinn R, et al. Exercise dependent modulation of immunological response pathways in endurance athletes with and without atrial fibrillation. *J Am Heart Assoc*. 2025;13:e033640.
21 DesOrmeaux GJ, Brunetta HS, Barbeau PA, et al. Excessive training does not induce mitochondrial dysfunction or impair insulin signalling within skeletal muscle. *J Physiol*. 2026;605 (Epub ahead of print) https://doi.org/10.1113/JP289538 (Epub ahead of print).
22 Nemkov T, Stauffer E, Cendali F, et al. Long-distance trail running induces inflammation-associated protein, lipid, and purine oxidation in red blood cells. *Blood R C I*. 2026;2:100055.
23 Srisawat K, Stead CA, Hesketh K, et al. People with obesity exhibit losses in muscle proteostasis that are partly improved by exercise training. *Proteomics*. 2024;24:2300395.
24 Viken KJ, Olsen T, Drevon CA, et al. Intra-individual variations in how insulin sensitivity responds to long-term exercise: predictions by machine learning based on large-scale serum proteomics. *Metabolites*. 2025;14:335.
25 Lee-Ødegård S, Hjorth M, Olsen T, et al. Serum proteomic profiling of physical activity reveals CD300LG as a novel exerkine with a potential causal link to glucose homeostasis. *Elife*. 2024;13:RP96535.
26 Zhu T, Hu X, Hwang J, et al. Urinary proteomics of simulated firefighting tasks and its relation to fitness parameters. *Int J Environ Res Public Health*. 2021;18:10618.
27 Cleland JGF, Ferreira JP, Mariottoni B, et al. The effect of spironolactone on cardiovascular function and markers of

fibrosis in people at increased risk of developing heart failure: the heart 'OMics' in AGEing (HOMAGE) randomized clinical trial. *Eur Heart J*. 2021;42:684–696.
28 Vlahou A, Hallinan D, Apweiler R, et al. Data sharing under the general data protection regulation: time to organize law and research ethics? *Hypertension*. 2021;77:1029–1035.
29 Yu YL, Siwy J, An DW, et al. Urinary proteomic signature of mineralocorticoid receptor antagonism by spironolactone: evidence from the HOMAGE trial. *Heart*. 2024;110:1180–1187.
30 Jantos-Siwy J, Schiffer E, Brand K, et al. Quantitative urinary proteome analysis for biomarker evaluation in chronic kidney disease. *J Proteome Res*. 2009;8:268–281.
31 Lazar C, Gatto L, Ferro M, Bruley C, Burger T. Accounting for the multiple natures of missing values in label-free quantitative proteomics data sets to compare imputation strategies. *J Proteome Res*. 2016;15:1116–1125.
32 Blom G. Statistical estimates and transformed beta-variables. *Biom J*. 1961;3:285.
33 An DW, Yu YL, Martens DS, et al. Statistical approaches applicable in managing OMICS data: urinary proteomics as exemplary case. *Mass Spec Rev*. 2023;42:1–18.
34 Stalmach A, Johnsson H, McInnes IB, et al. Identification of urinary peptide biomarkers associated with rheumatoid arthritis. *PLoS One*. 2014;9:e104625.
35 Levey AS, Stevens LA, Schmid CH, et al. A new equation to estimate glomerular filtration rate. *Ann Intern Med*. 2009;150:604–612.
36 SAS Institute Inc. The factor procedure. In: *SAS/STAT 131® User's Guide*. 2013. Cary, NC, USA.
37 Devereux RB, Alonso DR, Lutas EM, et al. Echocardiographic assessment of left ventricular hypertrophy: comparison to necropsy findings. *Am J Cardiol*. 1986;57:450–458.
38 Meinshausen N, Bühlmann P. Stability selection. *J Royal Stat Soc B*. 2010;72:417–473.
39 Lundberg SM, Lee SI. A unified approach to interpreting model predictions. In: *31st Conference on Neural Information Processing System (NIPS), Long Beach, CA, USA*; 2017. Internet communication. https://dl.acm.org/doi/10.5555/3295222.3295230. Accessed June 5, 2026
40 Yu G, He QY. ReactomePA: an R/Bioconductor package for reactome pathway analysis and visualization. *Mol BioSyst*. 2016;12:477–479.
41 Wu T, Hu E, Xu S, et al. clusterProfiler 4.0: a universal enrichment tool for interpreting omics data. *Innovation*. 2021;2:100141.
42 Bouchard C, Lortie G, Simoneau JA, Leblanc C, Thériault G, Tremblay A. Submaximal power output in adopted and biological siblings. *Ann Hum Biol*. 1984;11:303–309.
43 Takakura H, Furuichi Y, Yamada T, et al. Endurance training facilitates myoglobin desaturation during muscle contraction in rat skeletal muscle. *Sci Rep*. 2015;5:9403.
44 Xia XD, Zhou Z, Yu XH, Zheng XL, Tang CK. Myocardin: a novel player in atherosclerosis. *Atherosclerosis*. 2017;257:266–278.
45 Arystarkhova E, Ralph DL, Liu YB, Bouley R, McDonough AA, Sweadner KJ. Paradoxical activation of the sodium chloride cotransporter (NCC) without hypertension in kidney deficient in a regulatory subunit of Na,K-ATPase, FXYD2. *Physiol Rep*. 2014;2:e12226.
46 Oliva M, Muñoz-Aguirre M, Kim-Hellmuth S, et al. The impact of sex on gene expression across human tissues. *Science*. 2020;369:1331.
47 Granzier HL, Labeit S. Discovery of titin and it's role in heart function and disease. *Circ Res*. 2025;136:135–157.
48 Christensen AH, Pan G, Marvig RL, et al. Gain-of-function enhancer variant near KCNB1 causes familial ST-depression syndrome. *Eur Heart J*. 2025;46:3486–3497.
49 Lazarev VF, Guzhova I, Margulis BA. Glyceraldehyde-3-phosphate dehydrogenase is a multifaceted therapeutic target. *Pharmaceutics*. 2020;12:416.
50 Billings CG, Hurdman JA, Condliffe R, et al. Incremental shuttle walk test distance and autonomic dysfunction predict survival in pulmonary arterial hypertension. *J Heart Lung Transplant*. 2017;36:871–879.
51 Greenwood SA, Castle E, Lindup H, et al. Mortality and morbidity following exercise-based renal rehabilitation in patients with chronic kidney disease: the effect of programme completion and change in exercise capacity. *Nephrol Dial Transplant*. 2019;34:618–625.
52 Morales FJ, Montemayor T, Martinez A. Shuttle versus six-minute walk test in the prediction of outcome in chronic heart failure. *Int J Cardiol*. 2000;76:101–105.
53 Kuznetsova T, Herbots L, López B, et al. Prevalence of left ventricular diastolic dysfunction in a general population. *Circ Heart Fail*. 2009;2:105–112.
54 Rangaswami J, Bhalla V, Blair JEA, et al. Cardiorenal syndrome: classification, pathophysiology, diagnosis, and treatment strategies. *Circulation*. 2019;139:e840–e878.
55 Wilkinson TJ, Gould DW, Nixon DGD, Watson EL, Smith AC. Quality or quantity? Association of skeletal muscle myosteatosis and myofibrosis on physical function in chronic kidney disease. *Nephrol Dial Transpl*. 2018;34:1344–1353.