

THE LANCET

Healthy Longevity

Supplementary appendix

This appendix formed part of the original submission and has been peer reviewed. We post it as supplied by the authors.

Supplement to: Martens DS, Thijs L, Latosinska A, et al. Urinary peptidomic profiles to address age-related disabilities: a prospective population study. *Lancet Healthy Longev* 2021; published online Oct 14. [https://doi.org/10.1016/S2666-7568\(21\)00226-9](https://doi.org/10.1016/S2666-7568(21)00226-9).

The Lancet Healthy Longevity

Appendix

This web appendix formed part of the original submission and has been peer reviewed.
Supplement to: *Urinary peptidomic profiles to address age-related disabilities: a prospective population study*

DS Martens, L Thijs, A Latosinska, S Trenson, J Siwy, Z-Y Zhang, C Wang, J Beige, A Vlahou, S Janssens, H Mischak, TS Nawrot, JA Staessen, the FLEMENGHO investigators

Table of contents

Urinary proteomics	p2
Sample preparation and CE-MS analysis	p2
CE-MS data processing	p3
Sequencing of peptides	p3
References	p4
FLEMENGHO investigators	p5
Table 1 Use of medications in 778 FLEMENGHO examined in 2005-2010	p6
Table 2 Association of 210 sequenced urinary peptides with age in the 2005-2010 FLEMENGHO derivation dataset	p7
Table 3 Association of age with 39 sequenced urinary peptides with the highest significance per identified protein in the FLEMENGHO derivation and internal validation datasets	p19
Table 4 Number of replicated sequenced urinary peptides associated with age in the FLEMENGHO derivation and internal validation datasets	p21
Table 5 Weights assigned by elastic net regression to the multidimensional UPP biomarker reflecting ageing	p23
Table 6 Risk biomarkers in relation to age in 778 FLEMENGHO with adjustment for clustering within families	p27
Table 7 Risk of adverse health outcomes in relation to age in 778 FLEMENGHO with adjustment for clustering within families	p28
Table 8 Pathway analysis	p29
Table 9 Characteristics of FLEMENGHO participants analysed and not analysed	p32
Figure 1 FLEMENGHO timelines	p33
Figure 2 Analysis strategy	p34
Figure 3 Residual plot of UPP-age on C-age	p35
Figure 4 Short-term reproducibility of the UPP ageing biomarker	p36
Figure 5 Volcano plots for the FLEMENGHO datasets	p37

Urinary proteomics

Sample preparation and CE-MS analysis

Sample preparation and CE-MS analysis were performed essentially as described.¹ Urine aliquots were thawed and 700 μ l mixed with 700 μ l of 2 M urea, 10 mM NH_4OH containing 0.02 % SDS. Subsequently, samples were ultrafiltered using a Centrstat 20 kDa cut-off centrifugal filter device (Satorius, Göttingen, Germany) to eliminate high molecular weight proteins. The obtained filtrate was desalted using a PD 10 gel filtration column (GE Healthcare Bio Sciences, Uppsala, Sweden) to remove urea, electrolytes and salts as well as to enrich polypeptides. The samples were lyophilised and stored at 4°C before usage. Shortly before CE-MS analysis, the samples were re-suspended in 10 μ L HPLC-grade H_2O . Samples were injected into CE-MS with 2 psi for 99 sec, resulting in injection volumes of ~280 nL.

A P/ACE MDQ capillary electrophoresis system (Beckman Coulter, Fullerton, CA) was coupled with a Micro-TOF MS (Bruker Daltonic, Bremen, Germany). A solution of 20% acetonitrile (Sigma-Aldrich, Taufkirchen, Germany) in HPLC-grade water (Roth, Karlsruhe, Germany) supplemented with 0.94% formic acid (Sigma-Aldrich) was used as running buffer. For CE-MS analysis, the electrospray ionisation interface from Agilent Technologies (Palo Alto, CA) was set to a potential of -4.0 to -4.5 kV. Spectra were recorded over an m/z range of 350-3000 and accumulated every 3 s.

CE-MS data processing

After the CE-MS analysis, mass spectral ion peaks representing identical molecules at different charge states were deconvoluted into single masses using MosaFinder software.² Only signals with $z > 1$ observed in a minimum of 3 consecutive spectra with a signal-to-noise ratio of at least 4 were considered. The resulting peak list characterizes each polypeptide by

its mass and migration time. Data were calibrated utilising 150 reference mass data points and 452 reference migration time data points by applying global and local linear regression, respectively. Reference signals of 29 abundant collagen peptide fragments were used as internal standards for calibration using local linear regression.³ These 29 peptides were selected, based on their high abundance and lack of association with disease. This procedure reproducibly addresses both analytical and dilution variances in a single calibration step and performs better than standardisation based on the urinary creatinine concentration.³ The obtained peak list characterises each polypeptide by its calibrated molecular mass [Da], calibrated CE migration time [min] and normalised signal intensity. All detected peptides are deposited, matched, and annotated in a Microsoft SQL database allowing further statistical analysis.

Sequencing of peptides

Candidate biomarkers were sequenced using CE-MS/MS or LC-MS/MS analysis, as described in detail.⁴ MS/MS experiments were using an Ultimate 3000 nano-flow system (Dionex/LC Packings, USA) or a P/ACE MDQ capillary electrophoresis system (Beckman Coulter, Fullerton, CA), both connected to an LTQ Orbitrap hybrid mass spectrometer (Thermo Fisher Scientific, Germany) equipped with a nano-electrospray ion source. The mass spectrometer is operated in data-dependent mode to automatically switch between MS and MS/MS acquisition. Survey full-scan MS spectra (from m/z 300–2,000) were acquired in the Orbitrap. Ions were sequentially isolated for fragmentation. Data files were searched against the UniProt human nonredundant database using Proteome Discoverer 2.4 and the SEQUEST search engine. Relevant settings were: no fixed modifications, oxidation of methionine and proline as variable modifications. The minimum precursor mass was set to 790 Da, maximum precursor mass to 6000 Da with a minimum peak count of 10. The high-

confidence peptides were defined by cross-correlation (X_{corr}) >1.9 and rank = 1. Precursor mass tolerance and fragment mass tolerance were 5 ppm and 0.05 Da, respectively. For further validation of obtained peptide sequences, the correlation between peptide charge at the working pH of 2 and CE-migration time was utilised to minimise incorrect sequence assignment:⁵ calculated CE-migration time of the sequence candidate based on its peptide sequence (number of basic amino acids) was compared to the experimental migration time.

References

1. Mischak H, Vlahou A, Ioannidis JPA. Technical aspects and inter-laboratory variability in native peptide profiling: the CE-MS experience. *Clin Biochem* 2013; **46**: 432-43.
2. 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.
3. 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**: 268-81.
4. Klein J, Papadopoulos T, Mischak H, Mullen W. Comparison of CE-MS/MS and LC-MS/MS sequencing demonstrates significant complementarity in natural peptide identification in human urine. *Electrophoresis* 2014; **35**: 1060-64.
5. Zürgbig P, Renfrow MB, Schiffer E, et al. Biomarker discovery by CE-MS enables sequence analysis via MS/MS with platform-independent separation. *Electrophoresis* 2006; **27**: 2111-25.

FLEMENGHO investigators

Research Unit Hypertension and Cardiovascular Epidemiology, KU Leuven Department of Cardiovascular Sciences, University of Leuven, Leuven, Belgium — Prof K Asayama MD (current), Prof M Bochud MD (2008-2009), Prof J Boggia MD (2009-2012), Prof J Brguljan-Hitij MD (2011-2012), Prof YM Feng MD (2017-2019), YM Gu (2012-2016), A Hara PhD (2015), QF Huang MD (2016-2019), Y Jin MD (2008-2012), J Seidlerova MD (2008), YP Liu MD (2011-2014), JD Melgarejo MD (current), P Moliterno MSc (2017), Prof AN Odili MD (2010-2012), T Petit MD (2015-2016), A Raaijmakers MD (2015-2017), Prof R Schutte PhD (2012), Prof JA Staessen MD (current), Prof K Stolarz-Skrzypek MD (2009-2012), L Thijs MSc (current), V Tikhonoff MD (2008-2010), Prof JG Wang MD (2007-2008), FF Wei MD (2014-2019), DM Wei MD (current), Prof WY Yang MD (2014-2018), YL Yu MD (2019), Prof ZY Zhang MD (current).

Centre for Environmental Sciences, Hasselt University, Hasselt, Belgium — DS Martens PhD, Prof TS Nawrot PhD, HA Roels PhD, C Wang MSc.

Mosaiques-Diagnostics GmbH, Hannover, Germany — A Latosinska PhD, Prof H Mischak PhD, Justyna Siwy PhD.

Non-Profit Research Institute Alliance for the Promotion of Preventive Medicine, Mechelen Belgium — Prof K Asayama, Prof TW Hansen MD, Prof GE Maestre, Prof JA Staessen MD.

Biomedical Sciences Group, Faculty of Medicine, University of Leuven, Leuven, Belgium — Prof JA Staessen MD.

Table 1: Use of medications in 778 FLEMENGHO examined in 2005-2010

Medications (number taking drug [%])	C-age		UPP-age		UPP-age-R	
	Δ (95% CI)	p value	Δ (95% CI)	p value	Δ (95% CI)	p value
Antihypertensive drugs						
Any (198 [25.4])	17.9 (15.6, 20.1)	<0.0001	14.8 (12.6, 16.9)	<0.0001	0.92 (-0.33, 2.17)	0.15
Diuretics (78 [10.0])	19.4 (16.0, 22.8)	<0.0001	17.1 (14.0, 20.2)	<0.0001	2.07 (0.35, 3.79)	0.018
β Bs (114 [14.7])	17.5 (14.6, 20.3)	<0.0001	14.8 (12.2, 17.4)	<0.0001	1.27 (-0.20, 2.75)	0.091
RAS inhibitors including β Bs (21 [2.7])	22.6 (16.2, 28.9)	<0.0001	20.8 (15.1, 26.5)	<0.0001	3.34 (0.22, 6.46)	0.036
RAS inhibitors excluding β Bs (45 [5.8])	17.7 (13.3, 22.1)	<0.0001	13.4 (9.4, 17.4)	<0.0001	-0.33 (-2.62, 1.95)	0.78
Vasodilators (38 [4.9])	18.5 (13.8, 23.3)	<0.0001	13.7 (9.4, 17.9)	<0.0001	-0.69 (-3.02, 1.63)	0.56
Monotherapy (41 [5.3])	12.3 (7.8, 16.9)	<0.0001	12.2 (8.0, 16.4)	<0.0001	2.65 (0.32, 4.98)	0.026
Combination therapy (157 [20.2])	19.3 (16.8, 21.8)	<0.0001	15.4 (13.1, 17.8)	<0.0001	0.47 (-0.86, 1.80)	0.49
Antidiabetic drugs (22 [2.8])	19.0 (12.8, 25.3)	<0.0001	17.6 (12.0, 23.2)	<0.0001	2.86 (-0.17, 5.88)	0.064
NSAID (73 [9.4])	5.6 (1.9, 9.2)	0.0028	4.9 (1.7, 8.1)	0.0029	0.60 (-1.19, 2.38)	0.51
Lipid-lowering drugs						
Any (110 [14.1])	17.4 (14.6, 20.3)	<0.0001	14.2 (11.5, 16.8)	<0.0001	0.69 (-0.81, 2.18)	0.37
Statins (101 [13.0])	17.2 (14.3, 20.2)	<0.0001	13.7 (10.9, 16.4)	<0.0001	0.30 (-1.23, 1.84)	0.70
Any combination of the above (298 [38.3])	14.5 (12.4, 16.5)	<0.0001	12.2 (10.3, 14.1)	<0.0001	1.00 (-0.10, 2.10)	0.075

Abbreviations: β Bs, b-blockers; RAS, renin-angiotensin system; NSAID, non-steroidal anti-inflammatory drugs. RAS inhibitors include angiotensin-conversion inhibitors and angiotensin receptor blockers with or without β Bs. Lipid-lowering drugs include fibrates, ezetimibe, cholestyramine and statins. Δ is the age difference, given with 95% confidence interval, was obtained by subtracting non-users from users. In all comparisons, 480 participants not taking any of the listed drugs were the reference. Because of combination therapy, numbers are not additive. C-age and UPP-age refer to chronological age and age as predicted by the UPP. UPP-age-R refers to the residual of the regression of UPP-age on C-age and reflects accelerated ageing as predicted by the UPP-age, independent of C-age.

Table 2: Association of 210 sequenced urinary peptides with age in the 2005-2010 FLEMENGHO derivation dataset

Peptide ID number	Amino-acid sequence	Parental proteins identified from the amino-acid sequence		β (95% confidence interval)	p value
		Accession number	Name		
e08360	TGYTGKMcESSVNYcE	FAT4 (Q6V0I7)	protocadherin fat 4	0.34 (0.27, 0.40)	3.49 E-19
e12851	LDGAKGDAGPAGPKGEpG SpGENGApG	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.32 (-0.39, -0.25)	1.34 E-16
e09697	PGpPGPHGPPGpmGPHG LpGP	COL25A1 (Q9BXS0)	collagen type XXV alpha 1 chain	0.30 (0.23, 0.37)	4.08 E-14
e17856	ARGNDGARGSDGQPGPp GppGTAGFpGSpGAKGEV GP	COL3A1 (P02461)	collagen type III alpha 1 chain	-0.30 (-0.38, -0.23)	4.05 E-13
e05560	GSpGSpGPDGKTGPPGp	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.29 (-0.36, -0.22)	5.88 E-13
e20065	ARGNDGARGSDGQpGpp GPPGTAGFPpGSpGAKGEV GpAGSpGSNGApG	COL3A1 (P02461)	collagen type III alpha 1 chain	-0.28 (-0.36, -0.21)	9.64 E-12
e19885	ERGEQGPAGSpGFQGLpG pAGppGEAGKpGEQGVPG DLGAPGPSG	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.28 (-0.35, -0.20)	3.22 E-11
e16966	PpGESGREGApGAEGSpG RDGSPGAKGDRGETGP	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.27 (-0.34, -0.20)	4.31 E-11
e10072	GLpGTGGPpGENGKpGEp GPKG	COL3A1 (P02461)	collagen type III alpha 1 chain	-0.26 (-0.34, -0.19)	4.37 E-11
e03016	ApGDRGEPGPpGP	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.27 (-0.35, -0.20)	4.38 E-11
e04820	GVNVpSYGPpGPPG	COL19A1 (Q14993)	collagen type XIX alpha 1 chain	-0.27 (-0.35, -0.20)	6.71 E-11
e04945	SpGSPGPDGKTGPpGP	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.27 (-0.34, -0.19)	1.59 E-10
e12949	LDGAKGDAGPAGpKGEpG SpGENGApG	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.25 (-0.32, -0.18)	3.27 E-10
e12986	ADGQpGAKGEpGDAGAK GDAGPPGPAGP	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.25 (-0.32, -0.18)	6.39 E-10
e13707	GApQNGEPGGKGERGA pGEKGEpGppG	COL3A1 (P02461)	collagen type III alpha 1 chain	-0.25 (-0.32, -0.18)	3.31 E-09
e17922	NTGApGSPGVSGPKGDA GQpGEKGSpGAQpPGAP GPLG	COL3A1 (P02461)	collagen type III alpha 1 chain	-0.25 (-0.33, -0.18)	4.17 E-09

Peptide ID number	Amino-acid sequence	Parental proteins identified from the amino-acid sequence		β (95% confidence interval)	p value
		Accession number	Name		
e17212	RTGEVGA VGPPGFAGEKGS PSGEAGTAGPPTpGPQG	COL1A2 (P08123)	collagen type I alpha 2 chain	-0.24 (-0.31, -0.17)	4.86 E-09
e11213	ADGQPGAKGEPGDAGAK GDAGPPGP	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.24 (-0.31, -0.16)	7.21 E-09
e02286	PpGEAGKpGEQG	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.24 (-0.32, -0.17)	9.31 E-09
e11415	IGPpGPAGApGDKGESGP SGPAGPTG	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.24 (-0.31, -0.16)	9.55 E-09
e11073	NSGEpGApGSKGDTGAKG EPGpVG	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.24 (-0.32, -0.17)	1.17 E-08
e06441	TGSpGSpGPDGKTGPpGP	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.22 (-0.29, -0.15)	1.37 E-08
e17983	DGAPGQKGEMGPAGPTG PRGFpGppGPDGLPGSMG PP	COL4A1 (P02462)	collagen type IV alpha 1 chain	-0.25 (-0.32, -0.17)	1.38 E-08
e19745	SKGESGNKGEpGSAGPQ GPpGpSGEEGKRGNGEA GSAGPPGPpG	COL1A2 (P08123)	collagen type I alpha 2 chain	-0.24 (-0.31, -0.16)	1.39 E-08
e10987	VGEpGpAGSKGESGNKGE PGSAGP	COL1A2 (P08123)	collagen type I alpha 2 chain	-0.24 (-0.31, -0.16)	1.50 E-08
e04419	ApGDRGEpGPpGPAG	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.22 (-0.30, -0.15)	1.65 E-08
e06068	NDGApGKNGERGGpGGP	COL3A1 (P02461)	collagen type III alpha 1 chain	0.24 (0.16, 0.32)	2.09 E-08
e06316	GpPGEAGKpGEQGVpGD	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.23 (-0.31, -0.16)	3.54 E-08
e06733	AGSEADHEGTHSTKRG	FGA (P02671)	fibrinogen alpha chain	0.23 (0.16, 0.31)	5.44 E-08
e06709	GSpGSpGPDGKTGPpGPA G	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.22 (-0.29, -0.15)	6.94 E-08
e03025	GQDGRpGPpGPpG	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.22 (-0.30, -0.15)	7.34 E-08
e03322	SpGPDGKTGPpGPA	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.22 (-0.29, -0.15)	7.61 E-08
e06617	MpGSpGGpGSDGKpGPpG	COL3A1 (P02461)	collagen type III alpha 1 chain	-0.22 (-0.29, -0.15)	1.04 E-07
e17280	AAGEpGKAGERGVpGPpG AVGPAGKDGEAGAQQPP GP	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.22 (-0.30, -0.15)	1.47 E-07
e09891	SEGSPGHpGQPGpGPpG ApGP	COL3A1 (P02461)	collagen type III alpha 1 chain	-0.22 (-0.30, -0.15)	1.92 E-07

Peptide ID number	Amino-acid sequence	Parental proteins identified from the amino-acid sequence		β (95% confidence interval)	p value
		Accession number	Name		
e04321	PpGEAGKpGEQGVp	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.22 (-0.29, -0.14)	2.63 E-07
e13033	GpAGPPGEKGEPEGDDGP SGAEGPpGPQ	COL2A1 (P02458)	collagen type II alpha 1 chain	-0.22 (-0.29, -0.14)	3.05 E-07
e12508	KGNSEpGAPGSKGDTG AKGEpGpVG	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.22 (-0.29, -0.14)	3.26 E-07
e07678	TGSpGSpGPDGKTGPpGP AG	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.21 (-0.28, -0.14)	3.89 E-07
e08978	DEAGSEADHEGTHSTKRG	FGA (P02671)	fibrinogen alpha chain	0.22 (0.14, 0.29)	4.52 E-07
e02667	SpGPDGKTGpPGP	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.22 (-0.29, -0.14)	4.84 E-07
e15129	ERGEAGlpGVpGAKGEDG KDGSPGEpGANG	COL3A1 (P02461)	collagen type III alpha 1 chain	-0.21 (-0.28, -0.13)	7.19 E-07
e11013	AGpPGEAGKpGEQGVPG DLGApGP	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.22 (-0.29, -0.14)	8.12 E-07
e15778	ESGREGAPGAEGSPGRD GSpGAKGDRGETGp	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.21 (-0.28, -0.13)	8.14 E-07
e17288	GpGESGREGAPGAEGSp GRDGSpGAKGDRGETGp	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.21 (-0.29, -0.14)	9.74 E-07
e10025	NSGEpGApGSKGDTGAKG EpGP	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.21 (-0.29, -0.14)	9.74 E-07
e00966	YQTNKAKH	CSTB (P04080)	cystatin B	0.20 (0.13, 0.27)	1.02 E-06
e06961	GLpGTGGPpGENGKPEp	COL3A1 (P02461)	collagen type III alpha 1 chain	-0.21 (-0.28, -0.13)	1.09 E-06
e01399	DGRpGPpGpG	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.20 (-0.27, -0.13)	1.26 E-06
e17890	GpGADGQPGAkGEpGD AGAKGDAGPPGpAGPAG PPGpIG	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.21 (-0.28, -0.13)	1.75 E-06
e05114	SpGSpGPDGKTGPPGp	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.21 (-0.28, -0.13)	2.01 E-06
e11777	ANGApGNDGAKGDAGAp GApGSQGAPG	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.20 (-0.28, -0.13)	2.41 E-06
e17787	GARGLpGpPGSNGNPGPP GPSGSPGKDGPPGPAGN TGAP	COL3A1 (P02461)	collagen type III alpha 1 chain	-0.20 (-0.28, -0.13)	2.42 E-06

Peptide ID number	Amino-acid sequence	Parental proteins identified from the amino-acid sequence		β (95% confidence interval)	p value
		Accession number	Name		
e18864	SGPPGRAGEPGLQGPGAG PpGEKGEPEGDDGpSGAEG PpGPQG	COL2A1 (P02458)	collagen type II alpha 1 chain	-0.20 (-0.27, -0.12)	2.45 E-06
e04169	PpGEAGKpGEQGVp	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.20 (-0.27, -0.12)	2.52 E-06
e11753	GADGQPGAKGEPGDAGA KGDAGppGP	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.20 (-0.28, -0.13)	2.79 E-06
e10277	nGDDGEAGKpGRPGERG PPGp	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.20 (-0.28, -0.12)	2.79 E-06
e11902	ANGApGNDGAKGDAGAp GApGSQGApG	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.20 (-0.28, -0.12)	2.81 E-06
e09435	GEKGPSGEAGTAGPpGTp GPQG	COL1A2 (P08123)	collagen type I alpha 2 chain	-0.19 (-0.26, -0.12)	3.00 E-06
e14204	GNEGpSGPPGpAGSPGER GAAGSGGPIGpPG	COL11A2 (P13942)	collagen type XI alpha 2 chain	-0.20 (-0.28, -0.12)	3.11 E-06
e14315	NRGERGSEGSPPHpGQp GPPGpPGApGP	COL3A1 (P02461)	collagen type III alpha 1 chain	-0.19 (-0.26, -0.12)	4.53 E-06
e12156	GPAGpPGEKGEPEGDDGP SGAEGPPGp	COL2A1 (P02458)	collagen type II alpha 1 chain	-0.20 (-0.27, -0.12)	4.61 E-06
e15237	ERGEAGIpGVpGAKGEDG KDGSpGEpGANG	COL3A1 (P02461)	collagen type III alpha 1 chain	-0.19 (-0.27, -0.12)	5.61 E-06
e14015	PSTTVAPGSFSTAATTSPG ASGTTGVTTT	MUC19 (Q7Z5P9)	mucin 19, oligomeric	-0.19 (-0.27, -0.12)	7.24 E-06
e08188	GNDGApGKNGERGGpGG pGP	COL3A1 (P02461)	collagen type III alpha 1 chain	-0.19 (-0.26, -0.11)	7.58 E-06
e05259	DGQPGAKGEpGDAGAK	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.19 (-0.27, -0.12)	7.60 E-06
e16548	ADGQpGAKGEpGDAGAK GDAGPpGPAGPAGPPGPI G	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.19 (-0.26, -0.11)	7.65 E-06
e10266	NGDDGEAGKpGRpGERG PPGP	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.19 (-0.27, -0.12)	8.10 E-06
e18250	LALSKEKGDQGGPPGDPS PGSPGPAGPAGPPGYGp QGEP	COL4A3 (Q01955)	collagen type IV alpha 3 chain	-0.19 (-0.27, -0.11)	8.85 E-06
e03180	ApGDRGEpGPPGp	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.19 (-0.27, -0.11)	1.05 E-05

Peptide ID number	Amino-acid sequence	Parental proteins identified from the amino-acid sequence		β (95% confidence interval)	p value
		Accession number	Name		
e06154	EDGHpGKpGRpGERG	COL1A2 (P08123)	collagen type I alpha 2 chain	-0.19 (-0.26, -0.11)	1.11 E-05
e13825	LTGPIGppGPAGAPGDKG ESGPSGPAGPTG	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.18 (-0.26, -0.11)	1.14 E-05
e17946	PpGPAGFAGPPGADGQP GAKGEpGDAGAKGDAGP PGPAGP	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.19 (-0.26, -0.11)	1.47 E-05
e02311	GpPGPpGpGpPpA	COL8A1 (P27658)	collagen type VIII alpha 1 chain	-0.18 (-0.26, -0.11)	2.26 E-05
e07676	NDGAPGKNGERGGpGGp Gp	COL3A1 (P02461)	collagen type III alpha 1 chain	-0.18 (-0.26, -0.11)	2.49 E-05
e14837	ERGSpGPAGPKGSpGEAG RpGEAGLpGAKG	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.18 (-0.25, -0.10)	2.64 E-05
e17311	NTGApGSpGVSGpKGDAG QpGEKGSPPGAP Gp	COL3A1 (P02461)	collagen type III alpha 1 chain	-0.19 (-0.26, -0.11)	2.79 E-05
e02933	GppGPDGNKGEpG	COL1A2 (P08123)	collagen type I alpha 2 chain	-0.18 (-0.25, -0.10)	3.81 E-05
e06186	SpGSPGpDGKTGPPGpAG	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.18 (-0.25, -0.10)	3.99 E-05
e10386	DAGAPGApGGKGDAGAp GERGpPG	COL3A1 (P02461)	collagen type III alpha 1 chain	-0.17 (-0.25, -0.10)	4.99 E-05
e05551	VGPpGPPGpGpGPPS	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.18 (-0.25, -0.10)	5.59 E-05
e03638	SpGSpGPDGKTGPP	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.17 (-0.25, -0.10)	5.99 E-05
e06024	SpGSPGPDGKTGpPGPAG	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.17 (-0.25, -0.10)	6.47 E-05
e07807	GpPGpPGKNGDDGEAGKp G	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.17 (-0.24, -0.09)	6.49 E-05
e20087	ARGNDGARGSDGQpGpP GpGTAAGFpGSpGAKGEV GpAGSpGSNGApG	COL3A1 (P02461)	collagen type III alpha 1 chain	-0.17 (-0.25, -0.10)	6.94 E-05
e16540	GPAGAAGARGNDGQPPG AGPPGPVpAGGpGFpGA PG	COL2A1 (P02458)	collagen type II alpha 1 chain	-0.17 (-0.24, -0.09)	7.61 E-05
e16170	LTGSpGSpGpDGKTGPPG PAGQDGRPGpGppG	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.17 (-0.25, -0.09)	8.07 E-05
e11325	ADGQPGAKGEpGDAGAK GDAGPPGp	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.17 (-0.25, -0.09)	9.43 E-05

Peptide ID number	Amino-acid sequence	Parental proteins identified from the amino-acid sequence		β (95% confidence interval)	p value
		Accession number	Name		
e09408	GDDGEAGKpGRPGERGP PGp	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.17 (-0.24, -0.09)	9.43 E-05
e13792	KGGpGDQGEpGpQGLPGF SGPPGKEGE	COL9A2 (Q14055)	collagen type IX alpha 2 chain	-0.17 (-0.25, -0.09)	0.00010
e03248	EATEMcTQEc	PCDH9 (Q9HC56)	protocadherin 9	0.17 (0.09, 0.24)	0.00012
e16785	GSVGDPGMEGPMGQRG REGPMGpRGEAGpPGSG	COL17A1 (Q9UMD9)	collagen type XVII alpha 1 chain	-0.17 (-0.24, -0.09)	0.00013
e18041	NTGAPGSpGVSGpKGDAG QpGEKGSpGAQpGpGAPG PLG	COL3A1 (P02461)	collagen type III alpha 1 chain	-0.17 (-0.25, -0.09)	0.00013
e16219	ppGAKGQEGAHGAPGAA GNPpGAPGHVpGAPGpSGp P	COL22A1 (Q8NFW1)	collagen type XXII alpha 1 chain	-0.17 (-0.24, -0.09)	0.00014
e07132	EpGSpGENGAPGmGPR	COL1A1 (P02452)	collagen type I alpha 1 chain	0.16 (0.09, 0.24)	0.00015
e10771	DGQPGAKGEPGDAGAKG DAGPpGP	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.16 (-0.24, -0.09)	0.00015
e02795	GPpGPDGNKGEpG	COL1A2 (P08123)	collagen type I alpha 2 chain	-0.17 (-0.24, -0.09)	0.00016
e11008	pPGEEGKRGRGDpGTVG PpGP	COL5A2 (P05997)	collagen type V alpha 2 chain	-0.16 (-0.24, -0.09)	0.00016
e20509	QGPPGKNGETGPQGpG pTGpGGDKGDTGPPGPQ GLQGLPGTGGPPGENGK	COL3A1 (P02461)	collagen type III alpha 1 chain	-0.17 (-0.24, -0.09)	0.00016
e05074	DGHpGKpGRPGERG	COL1A2 (P08123)	collagen type I alpha 2 chain	-0.16 (-0.23, -0.08)	0.00019
e05830	GPpGVPGpPGpGGSPGLP	COL22A1 (Q8NFW1)	collagen type XXII alpha 1 chain	-0.16 (-0.24, -0.09)	0.00021
e03183	KPWAAQDGPKPG	LRRC25 (Q8N386)	leucine rich repeat containing 25	-0.15 (-0.23, -0.08)	0.00024
e08463	DEAGSEADHEGTHSTKR	FGA (P02671)	fibrinogen alpha chain	0.16 (0.08, 0.24)	0.00025
e10573	GPpGEAGKpGEQGVpGDL GAPGP	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.16 (-0.24, -0.09)	0.00025
e00467	DGKTGPpGPA	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.16 (-0.23, -0.08)	0.00026
e04843	GPpGpPGpGpPGPPS	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.16 (-0.23, -0.08)	0.00026
e11972	ADGQpGAKGEPGDAGAK GDAGpGPA	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.16 (-0.23, -0.08)	0.00026

Peptide ID number	Amino-acid sequence	Parental proteins identified from the amino-acid sequence		β (95% confidence interval)	p value
		Accession number	Name		
e01155	ApGDKGESGPS	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.16 (-0.23, -0.08)	0.00034
e12366	GANGApGNDGAKGDAGA pGApGSQGApG	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.15 (-0.23, -0.08)	0.00036
e10502	DAGApGApGGKGDAGApG ERGPpG	COL3A1 (P02461)	collagen type III alpha 1 chain	-0.16 (-0.23, -0.08)	0.00037
e18492	DLADGVSGGEGKGGSDG GGSHRKEGEEADAPGVIP GIVG	CD99 (P14209)	CD99 molecule (Xg blood group)	0.15 (0.07, 0.22)	0.00044
e06220	GHPDTLNQGEFKEL	S100A9 (P06702)	S100 calcium binding protein A9	-0.15 (-0.23, -0.08)	0.00047
e03354	SpGERGETGPpGP	COL3A1 (P02461)	collagen type III alpha 1 chain	-0.15 (-0.23, -0.07)	0.00055
e13472	LRGGAGpPGPEGKGA GPpGPpGAAGTpG	COL3A1 (P02461)	collagen type III alpha 1 chain	-0.15 (-0.22, -0.07)	0.00058
e14735	KNGETGPQGGpTGPG GDKGDTGPpGPQG	COL3A1 (P02461)	collagen type III alpha 1 chain	-0.15 (-0.22, -0.07)	0.00058
e18377	FAEEKAVADTRDQADGSR ASVDSGSSEEQGGSSRA	PIGR (P01833)	polymeric immunoglobulin receptor	0.15 (0.07, 0.23)	0.00058
e18782	GRPEAQPPPLSSEHKPV AGDAVPgPKDGSAPVVR GA	VGF (O15240)	VGF nerve growth factor inducible	-0.15 (-0.22, -0.07)	0.00058
e08228	GEpGSpGENGApGQMGP RG	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.15 (-0.23, -0.08)	0.00063
e04947	EpGKAGERGVpGPpG	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.14 (-0.22, -0.07)	0.00066
e12714	pAGPPGEKGEpGDDGPS GAEGPpGPQ	COL2A1 (P02458)	collagen type II alpha 1 chain	-0.15 (-0.23, -0.07)	0.00069
e16656	FAGpPGADGQpGAKGEQ GEAGQKGDAGApGpQGP	COL2A1 (P02458)	collagen type II alpha 1 chain	-0.15 (-0.23, -0.07)	0.00081
e12235	GQNGEpGGKGERGApGE KGEgGPpG	COL3A1 (P02461)	collagen type III alpha 1 chain	-0.15 (-0.22, -0.07)	0.00083
e02716	KNDQnTSVSHA	BRWD3 (Q6RI45)	bromodomain and WD repeat domain containing 3	-0.15 (-0.22, -0.07)	0.00085
e10141	PpGEAGKpGEQGVpGDLG ApGP	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.15 (-0.23, -0.07)	0.00085

Peptide ID number	Amino-acid sequence	Parental proteins identified from the amino-acid sequence		β (95% confidence interval)	p value
		Accession number	Name		
e11780	QNGEpGGKGERGApGEK GEGGPpG	COL3A1 (P02461)	collagen type III alpha 1 chain	-0.15 (-0.23, -0.07)	0.00085
e06650	EGSpGRDGSpGAKGDRG	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.14 (-0.21, -0.06)	0.00086
e06584	DGApGKNGERGGpGGpG P	COL3A1 (P02461)	collagen type III alpha 1 chain	-0.15 (-0.23, -0.07)	0.00087
e13697	ApGPAGSRGApGPQGpRG DKGETGERG	COL3A1 (P02461)	collagen type III alpha 1 chain	-0.15 (-0.22, -0.07)	0.00089
e12556	KNGDDGEAGKpGRPGER GpPGPQG	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.14 (-0.22, -0.07)	0.00094
e13253	EDGKDGSpGEpGANGLpG AAGERGApG	COL3A1 (P02461)	collagen type III alpha 1 chain	-0.14 (-0.22, -0.07)	0.00094
e16417	RTGEVGA VGpPGFAGEK PSGEAGTAGPpGTpGP	COL1A2 (P08123)	collagen type I alpha 2 chain	-0.15 (-0.23, -0.07)	0.00095
e17596	GPpGESGREGApGAEGSp GRDGSPGAKGDRGETGp A	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.14 (-0.21, -0.07)	0.00096
e01923	SpGERGETGpP	COL3A1 (P02461)	collagen type III alpha 1 chain	-0.15 (-0.22, -0.07)	0.00098
e01758	ApGDRGEpGpP	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.14 (-0.22, -0.07)	0.0010
e00421	DpGKNGDKG	COL1A2 (P08123)	collagen type I alpha 2 chain	-0.15 (-0.22, -0.07)	0.0011
e09521	EGGKGAAGpPGPpGAAGT pGLQG	COL3A1 (P02461)	collagen type III alpha 1 chain	-0.14 (-0.22, -0.07)	0.0012
e13850	PpGKNGDDGEAGKpGRp GERGppGPQ	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.13 (-0.21, -0.06)	0.0015
e03890	ApGDRGEpGpPGPA	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.13 (-0.21, -0.06)	0.0017
e11002	AEGSpGRDGSpGAKGDR GETGPA	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.13 (-0.21, -0.06)	0.0017
e16419	EAGRDNpGNDGppGRD GQPGHKGERGypG	COL1A2 (P08123)	collagen type I alpha 2 chain	-0.13 (-0.21, -0.06)	0.0017
e02711	MNIGKpNAKDI	EP400 (Q96L91)	E1A binding protein P400	0.14 (0.06, 0.21)	0.0018
e15220	LRGGAGPpGPEGKGA GpPGpPGAAGTpGLQG	COL3A1 (P02461)	collagen type III alpha 1 chain	-0.14 (-0.21, -0.06)	0.0018

Peptide ID number	Amino-acid sequence	Parental proteins identified from the amino-acid sequence		β (95% confidence interval)	p value
		Accession number	Name		
e11536	NGDDGEAGKPRpGERG PpGPQG	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.13 (-0.20, -0.06)	0.0019
e05236	SpGSpGPDGKTGppGP	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.14 (-0.21, -0.06)	0.0020
e15323	VQGpGpAGEEGKRGAR GEPGPTGLpGpG	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.13 (-0.20, -0.06)	0.0021
e04979	GLpGTGGppGENGKPG	COL3A1 (P02461)	collagen type III alpha 1 chain	-0.14 (-0.21, -0.06)	0.0022
e16606	ADGQpGAKGEpGDAGAK GDAGpPGPAGPAGPpGPI G	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.13 (-0.21, -0.06)	0.0022
e08755	EGSpGRDGSpGAKGDRG ET	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.13 (-0.21, -0.06)	0.0023
e19868	EEKAVADTRDQADGSRAS VDSGSSEEQGSSRALVS TLVPL	PIGR (P01833)	polymeric immunoglobulin receptor	0.13 (0.06, 0.21)	0.0023
e04457	GSpGGpGSDGKpGpG	COL3A1 (P02461)	collagen type III alpha 1 chain	-0.13 (-0.21, -0.06)	0.0024
e13816	ERGEAGIpGvPpGAKGEDG KDGSpGEpG	COL3A1 (P02461)	collagen type III alpha 1 chain	-0.13 (-0.21, -0.06)	0.0025
e06922	GpGFGFRpGDpGPPGPPG	COL15A1 (P39059)	collagen type XV alpha 1 chain	-0.13 (-0.21, -0.06)	0.0027
e10326	SGEpGApGSKGDTGAKGE pGPVG	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.13 (-0.21, -0.06)	0.0029
e17690	pPGADGQPGAKGEpGDA GAKGDAGppGPAGPAGPP GPIG	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.13 (-0.21, -0.05)	0.0029
e04664	KSqAqSTDSSkDP	KRT12 (Q99456)	keratin 12	0.13 (0.05, 0.21)	0.0032
e15070	YGGSSGGSSGGGYGG GSSSGHGGSSSGGYGG GS	KRT10 (P13645)	keratin 10	0.13 (0.05, 0.21)	0.0032
e18044	GPTGAVGFAGPQpDGQ pGVKGEpGEPGQKGDAG SPGP	COL5A2 (P05997)	collagen type V alpha 2 chain	-0.13 (-0.21, -0.05)	0.0032
e18831	DKGETGEQGDRIKGRH GFSLQGppGppGSPGEQ GP	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.13 (-0.20, -0.05)	0.0040

Peptide ID number	Amino-acid sequence	Parental proteins identified from the amino-acid sequence		β (95% confidence interval)	p value
		Accession number	Name		
e01784	DGGGSPKGDVDP	FXYD2 (P54710)	FXYD domain containing ion transport regulator 2	-0.12 (-0.20, -0.05)	0.0045
e15467	GRDGNpGNDGPPGRDGGQpGHKGERGYpG	COL1A2 (P08123)	collagen type I alpha 2 chain	-0.12 (-0.20, -0.05)	0.0049
e10668	GpPGEAGKpGEEQVpGDLGAPGp	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.13 (-0.21, -0.05)	0.0050
e05369	DGQpGAKGepGDAGAK	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.12 (-0.20, -0.05)	0.0052
e05899	SAPYPGGGGNGASGGGSGGS	PCDH7 (O60245)	protocadherin 7	-0.13 (-0.20, -0.05)	0.0052
e01132	cDDYRLcE	MGP (P08493)	matrix Gla protein	0.12 (0.04, 0.19)	0.0053
e06309	VGPpGPpGPpGPpGPPSA	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.12 (-0.20, -0.05)	0.0054
e17254	ENGKpGEpGpKGDAGApGApGGKGDAGApGERGpPG	COL3A1 (P02461)	collagen type III alpha 1 chain	-0.12 (-0.20, -0.05)	0.0054
e11498	GNSGepGApGSKGDTGAKGEPGpVG	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.12 (-0.20, -0.05)	0.0055
e10863	DGQPpGAKGepGDAGAKGDAGPPGp	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.13 (-0.20, -0.05)	0.0056
e12101	ADGQpGAKGepGDAGAKGDAGPpGpA	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.12 (-0.20, -0.04)	0.0056
e09874	DGESGRpGRPGERGLpGPpG	COL3A1 (P02461)	collagen type III alpha 1 chain	-0.12 (-0.20, -0.05)	0.0060
e11375	NGApGNDGAKGDAGApGApGSQGApG	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.12 (-0.20, -0.05)	0.0060
e11537	GKNGDDGEAGKpGRpGERGpG	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.12 (-0.20, -0.04)	0.0063
e10483	GPpGPpGKNGDDGEAGKpGRpG	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.12 (-0.19, -0.04)	0.0068
e07225	pPGEAGKpGEEQVpGDLG	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.12 (-0.20, -0.04)	0.0069
e10518	DGPpGRDGGQpGHKGERGYpG	COL1A2 (P08123)	collagen type I alpha 2 chain	-0.12 (-0.20, -0.04)	0.0070
e17565	GEPGITGLPGEKGMAGGHGPDGPKGSpGpSGTpGDTG	COL5A2 (P05997)	collagen type V alpha 2 chain	0.12 (0.04, 0.19)	0.0071

Peptide ID number	Amino-acid sequence	Parental proteins identified from the amino-acid sequence		β (95% confidence interval)	p value
		Accession number	Name		
e11122	AGPpGEAGKpGEQGVpGD LGApGP	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.12 (-0.20, -0.04)	0.0075
e05671	KpGEQGVpGDLGAPGp	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.12 (-0.20, -0.04)	0.0076
e10445	GNSGEpGApGSKGDTGAK GEPGp	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.12 (-0.20, -0.04)	0.0076
e11193	NSGEpGApGSKGDTGAKG EpGPVG	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.12 (-0.20, -0.04)	0.0081
e05288	FPGQTGPRGEMGQp	COL7A1 (Q02388)	collagen type VII alpha 1 chain	0.12 (0.04, 0.20)	0.0084
e09449	KGNDGApGKNGERGGpG GpGp	COL3A1 (P02461)	collagen type III alpha 1 chain	-0.11 (-0.19, -0.04)	0.0084
e02592	DDGEAGKpGRpG	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.11 (-0.19, -0.04)	0.0085
e17047	PpGESGREGApGAEGSpG RDGSpGAKGDRGETGP	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.12 (-0.20, -0.04)	0.0085
e10881	NGEpGGKGERGAPGEKG EGGppG	COL3A1 (P02461)	collagen type III alpha 1 chain	-0.12 (-0.19, -0.04)	0.0088
e11452	ADGQpGAKGEpGDAGAK GDAGPPGp	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.12 (-0.20, -0.04)	0.0088
e11559	GRTGDAGPVGPPGPpGpp GpPGPPS	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.11 (-0.19, -0.04)	0.0090
e10032	EGSpGRDGSpGAKGDRG ETGP	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.12 (-0.20, -0.04)	0.0091
e02858	EAGGGNSLQNSP	FRMD4A (Q9P2Q2)	FERM domain containing 4A	-0.11 (-0.19, -0.04)	0.0094
e07275	NGApGNDGAKGDAGApG ApG	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.12 (-0.20, -0.04)	0.0094
e06839	GEDGHpGKpGRpGERG	COL1A2 (P08123)	collagen type I alpha 2 chain	-0.11 (-0.19, -0.04)	0.0096
e19279	TSPSQQLIPVKNLSENIEIL	PKD1L2 (Q7Z442)	polycystic kidney disease 1-like 2	-0.11 (-0.19, -0.04)	0.0099
e17095	NKTLIYIL	UBAC2 (Q8NBM4)	ubiquitin-associated domain-containing protein 2	0.11 (0.03, 0.18)	0.011
e02105	PIQEGAPGRPG	COL4A2 (P08572)	collagen type IV alpha 2 chain	-0.11 (-0.18, -0.03)	0.013
e17751	pPGPQGPSGLSIQGMpGm pGEKGEKGDGLPGPQG	COL14A1 (Q05707)	collagen type XIV alpha 1 chain	-0.11 (-0.19, -0.03)	0.015

Peptide ID number	Amino-acid sequence	Parental proteins identified from the amino-acid sequence		β (95% confidence interval)	p value
		Accession number	Name		
e08916	DDGEAGKpGRpGERGpP GP	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.11 (-0.18, -0.03)	0.016
e11850	ADGQPGAKGEpGDAGAK GDAGPPGpA	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.10 (-0.18, -0.03)	0.018
e17505	AGRpGEVGPpGPpGPAGE KGSPGADGPAGAPGTpGP QG	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.10 (-0.18, -0.03)	0.018
e17599	EEKAVADTRDQADGSRAS VDSGSSEEQGGSSRA	PIGR (P01833)	polymeric immunoglobulin receptor	0.10 (0.03, 0.18)	0.022
e00340	cDDYRLc	MGP (P08493)	matrix Gla protein	0.10 (0.02, 0.17)	0.024
e05353	GPpGKNGDDGEAGkPpG	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.10 (-0.18, -0.02)	0.024
e11906	QNGEpGGKGERGApGEK GEGGppG	COL3A1 (P02461)	collagen type III alpha 1 chain	-0.10 (-0.18, -0.02)	0.027
e07799	GPpGEAGKpGEQGVPGDL G	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.10 (-0.17, -0.02)	0.028
e12661	GKNGDDGEAGKpGRpGE RGpPpGpQ	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.10 (-0.17, -0.02)	0.031
e11668	GGpGSDGKPGppGSQGE SGRPGpPpG	COL3A1 (P02461)	collagen type III alpha 1 chain	-0.10 (-0.17, -0.02)	0.038
e06256	TGLSMDGGGSPKGDVDP	FXD2 (P54710)	FXD domain containing ion transport regulator 2	-0.09 (-0.17, -0.02)	0.039
e10554	EGSpGRDGSpGAKGDRG ETGPA	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.09 (-0.17, -0.01)	0.044
e00579	EGRQKESV	PPP1R3A (Q16821)	protein phosphatase 1 regulatory subunit 3A	0.09 (0.01, 0.17)	0.048
e04647	DGPpGRDGGpGHKG	COL1A2 (P08123)	collagen type I alpha 2 chain	-0.09 (-0.17, -0.01)	0.049
e16771	GADGQpGAKGEpGDAGA KGDAGPPGPAGPpG PIG	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.09 (-0.17, -0.01)	0.049

The derivation study included the baseline data of 559 FLEMENGHO participants. The lower-case letters in the amino-acid sequence identify post-translational modifications: c, disulphide bridges; m, oxidized methionine; n, deaminated asparagine; p, hydroxylated proline; k, hydroxylated lysine; q, deaminated glutamine. Accession numbers refer to the Uniprot database (<http://www.uniprot.org/uniprot>). Association sizes (β) between urinary peptides and age were expressed per 10-year increment and were adjusted for sex, mean arterial pressure (diastolic blood pressure plus one third of pulse pressure (the difference between systolic and diastolic blood pressure), body mass index, plasma glucose, γ -glutamyltransferase as index of alcohol intake, current smoking, the total-to-high-density-lipoprotein serum cholesterol ratio, and the glomerular filtration rate calculated from serum creatinine by the Chronic Kidney Disease Epidemiology Collaboration equation. p-values were corrected for the Benjamini-Hochberg false discovery rate for 635 tested associations.

Table 3: Association of age with 39 sequenced urinary peptides with the highest significance per identified protein in the FLEMENGHO derivation and internal validation studies

Peptide ID number	Parental protein	Derivation and time-shifted internal validation study (n=559)				Synchronous internal validation study (n=219) 2005-2010	
		2005-2010 (baseline)		2009-2013 (follow-up)		β (95% CI)	p value
		β (95% CI)	p value	β (95% CI)	p value		
e08360	FAT4*	0.34 (0.27, 0.40)	1.09 E-21	0.25 (0.18, 0.32)	7.98 E-11	0.34 (0.24, 0.44)	1.01 E-10
e12851	COL1A1*	-0.32 (-0.39, -0.25)	6.34 E-19	-0.19 (-0.26, -0.11)	8.61 E-07	-0.34 (-0.44, -0.23)	2.36 E-09
e09697	COL25A1*	0.30 (0.23, 0.37)	3.21 E-16	0.37 (0.30, 0.44)	3.91 E-23	0.25 (0.14, 0.36)	7.07 E-06
e17856	COL3A1*	-0.30 (-0.38, -0.23)	3.83 E-15	-0.17 (-0.24, -0.09)	2.75 E-05	-0.34 (-0.45, -0.22)	1.39 E-08
e04820	COL19A1*	-0.27 (-0.35, -0.20)	1.37 E-12	-0.15 (-0.23, -0.08)	8.17 E-05	-0.21 (-0.32, -0.09)	0.00037
e17212	COL1A2*	-0.24 (-0.31, -0.17)	1.53 E-10	-0.21 (-0.29, -0.14)	5.07 E-08	-0.25 (-0.36, -0.14)	9.62 E-06
e17983	COL4A1*	-0.25 (-0.32, -0.17)	5.44 E-10	-0.18 (-0.25, -0.10)	1.09 E-05	-0.18 (-0.30, -0.06)	0.0029
e06733	FGA*	0.23 (0.16, 0.31)	2.74 E-09	0.08 (0.00, 0.16)	0.054	0.18 (0.06, 0.30)	0.0031
e13033	COL2A1*	-0.22 (-0.29, -0.14)	1.97 E-08	-0.08 (-0.15, 0.00)	0.056	-0.17 (-0.29, -0.05)	0.0051
e00966	CSTB*	0.20 (0.13, 0.27)	8.68 E-08	0.17 (0.10, 0.25)	1.61 E-05	0.20 (0.08, 0.32)	0.00082
e14204	COL11A2*	-0.20 (-0.28, -0.12)	3.38 E-07	-0.13 (-0.21, -0.06)	0.00079	-0.22 (-0.34, -0.11)	0.00017
e14015	MUC19*	-0.19 (-0.27, -0.12)	8.32 E-07	-0.10 (-0.17, -0.02)	0.017	-0.21 (-0.33, -0.09)	0.00054
e18250	COL4A3*	-0.19 (-0.27, -0.11)	1.09 E-06	-0.18 (-0.26, -0.11)	3.37 E-06	-0.18 (-0.30, -0.06)	0.0035
e02311	COL8A1*	-0.18 (-0.26, -0.11)	2.95 E-06	-0.09 (-0.17, -0.01)	0.022	-0.22 (-0.33, -0.10)	0.00019
e13792	COL9A2*	-0.17 (-0.25, -0.09)	1.64 E-05	-0.08 (-0.16, 0.00)	0.057	-0.20 (-0.33, -0.08)	0.0014
e03248	PCDH9*	0.17 (0.09, 0.24)	2.01 E-05	0.13 (0.06, 0.21)	0.00071	0.18 (0.07, 0.30)	0.0022
e16785	COL17A1*	-0.17 (-0.24, -0.09)	2.15 E-05	-0.08 (-0.15, 0.00)	0.058	-0.17 (-0.29, -0.05)	0.0065
e16219	COL22A1*	-0.17 (-0.24, -0.09)	2.27 E-05	-0.08 (-0.16, 0.00)	0.047	-0.25 (-0.36, -0.14)	2.31 E-05
e11008	COL5A2*	-0.16 (-0.24, -0.09)	2.85 E-05	-0.07 (-0.15, 0.00)	0.059	-0.19 (-0.31, -0.07)	0.0016
e03183	LRRRC25*	-0.15 (-0.23, -0.08)	4.47 E-05	-0.08 (-0.16, -0.01)	0.036	-0.19 (-0.30, -0.08)	0.00078
e18492	CD99§	0.15 (0.07, 0.22)	8.98 E-05	0.04 (-0.04, 0.12)	0.32	0.00 (-0.11, 0.12)	0.97

Peptide ID number	Parental protein	Derivation and time-shifted internal validation study (n=559)				Synchronous internal validation study (n=219) 2005-2010	
		2005-2010 (baseline)		2009-2013 (follow-up)		β (95% CI)	p value
		β (95% CI)	p value	β (95% CI)	p value		
e06220	S100A9*	-0.15 (-0.23, -0.08)	9.59 E-05	-0.07 (-0.15, 0.01)	0.079	-0.14 (-0.26, -0.02)	0.026
e18377	PIGR§	0.15 (0.07, 0.23)	0.00012	-0.04 (-0.12, 0.04)	0.36	-0.02 (-0.14, 0.11)	0.80
e18782	VEGF*	-0.15 (-0.22, -0.07)	0.00012	-0.10 (-0.18, -0.03)	0.0086	-0.14 (-0.26, -0.02)	0.023
e02716	BRWD3*	-0.15 (-0.22, -0.07)	0.00019	-0.07 (-0.15, 0.01)	0.081	-0.12 (-0.24, 0.00)	0.059
e02711	EP400‡	0.14 (0.06, 0.21)	0.00047	-0.03 (-0.11, 0.05)	0.49	0.12 (0.00, 0.24)	0.049
e06922	COL15A1‡	-0.13 (-0.21, -0.06)	0.00077	-0.06 (-0.14, 0.02)	0.12	-0.21 (-0.33, -0.09)	0.00073
e04664	KRT12†	0.13 (0.05, 0.21)	0.00092	0.10 (0.02, 0.17)	0.019	0.08 (-0.04, 0.19)	0.19
e15070	KRT10§	0.13 (0.05, 0.21)	0.00095	0.02 (-0.06, 0.10)	0.69	0.06 (-0.07, 0.18)	0.37
e01784	FXD2‡	-0.12 (-0.20, -0.05)	0.0013	-0.06 (-0.14, 0.02)	0.14	-0.11 (-0.23, 0.01)	0.080
e05899	PCDH7§	-0.13 (-0.20, -0.05)	0.0016	0.01 (-0.07, 0.08)	0.89	-0.05 (-0.18, 0.07)	0.37
e01132	MGP*	0.12 (0.04, 0.19)	0.0017	0.14 (0.06, 0.21)	0.00044	0.11 (-0.01, 0.22)	0.067
e05288	COL7A1§	0.12 (0.04, 0.20)	0.0031	0.05 (-0.03, 0.13)	0.20	0.07 (-0.05, 0.19)	0.26
e02858	FRMD4A‡	-0.11 (-0.19, -0.04)	0.0037	-0.04 (-0.12, 0.04)	0.35	-0.13 (-0.25, -0.01)	0.034
e19279	PK1L2*	-0.11 (-0.19, -0.04)	0.0038	-0.10 (-0.17, -0.02)	0.0097	-0.12 (-0.24, 0.01)	0.061
e17095	UBAC2§	0.11 (0.03, 0.18)	0.0043	-0.03 (-0.10, 0.05)	0.49	0.08 (-0.04, 0.20)	0.22
e02105	COL4A2‡	-0.11 (-0.18, -0.03)	0.0051	0.00 (-0.07, 0.08)	0.93	-0.15 (-0.27, -0.03)	0.012
e17751	COL14A1†	-0.11 (-0.19, -0.03)	0.0061	-0.09 (-0.17, -0.02)	0.018	-0.02 (-0.14, 0.10)	0.75
e00579	PPP1R3A‡	0.09 (0.01, 0.17)	0.022	-0.05 (-0.13, 0.03)	0.22	0.12 (-0.01, 0.24)	0.066

Urinary peptides with the highest significance per identified protein were selected for this analysis to demonstrate consistency of the associations in 559 participants included in the derivation (2005-2010) and the time-shifted internal validation (2009-2013) datasets and in the 219 participants included in the synchronous internal validation dataset (2005-2010). Association sizes (β) between the urinary peptides and age were expressed per 10-year increment and were adjusted for sex, mean arterial pressure (diastolic blood pressure plus one third of pulse pressure (the difference between systolic and diastolic blood pressure)), body mass index, plasma glucose, γ -glutamyltransferase as index of alcohol intake, current smoking, the total-to-high-density-lipoprotein serum cholesterol ratio, and the glomerular filtration rate calculated from serum creatinine by the Chronic Kidney Disease Epidemiology Collaboration equation. Symbols indicate validation status based on an p value of <0.10 (given the prior probability established by the significance in the derivation dataset: * replicated in the time-shifted and synchronous validation datasets; † replicated in the time-shifted internal validation dataset; ‡ replicated in the synchronous internal validation dataset; § not replicated).

Table 4: Number of replicated sequenced urinary peptides identifying parental proteins in the FLEMENGHO derivation and internal validation datasets

Peptide ID number	Parental protein	Number of peptide fragments (percentage replicated compared with derivation dataset)		
		Derivation dataset (2005-2010)	Time-shifted dataset (2009-2013)	Synchronous dataset (2005-2010)
e08360	FAT4	1	1 (100%)	1 (100%)
e12851	COL1A1	104	87 (79%)	94 (90%)
e09697	COL25A1	1	1 (100%)	1 (100%)
e17856	COL3A1	41	34 (83%)	39 (95%)
e04820	COL19A1	1	1 (100%)	1 (100%)
e17212	COL1A2	15	10 (67%)	13 (87%)
e17983	COL4A1	1	1 (100%)	1 (100%)
e06733	FGA	3	3 (100%)	3 (100%)
e13033	COL2A1	6	6 (100%)	5 (83%)
e00966	CSTB	1	1 (100%)	1 (100%)
e14204	COL11A2	1	1 (100%)	1 (100%)
e14015	MUC19	1	1 (100%)	1 (100%)
e18250	COL4A3	1	1 (100%)	1 (100%)
e02311	COL8A1	1	1 (100%)	1 (100%)
e13792	COL9A2	1	1 (100%)	1 (100%)
e03248	PCDH9	1	1 (100%)	1 (100%)
e16785	COL17A1	1	1 (100%)	1 (100%)
e16219	COL22A1	2	2 (100%)	2 (100%)
e11008	COL5A2	3	1 (33%)	2 (67%)
e03183	LRRC25	1	1 (100%)	1 (100%)
e18492	CD99	1	0 (0%)	0 (0%)
e06220	S100A9	1	1 (100%)	1 (100%)
e18377	PIGR	3	0 (0%)	0 (0%)
e18782	VGFB	1	1 (100%)	1 (100%)
e02716	BRWD3	1	1 (100%)	1 (100%)
e02711	EP400	1	0 (0%)	1 (100%)
e06922	COL15A1	1	0 (0%)	1 (100%)
e04664	KRT12	1	1 (100%)	0 (0%)
e15070	KRT10	1	0 (0%)	0 (0%)
e01784	FXYD2	2	1 (50%)	1 (50%)
e05899	PCDH7	1	0 (0%)	0 (0%)

Peptide ID number	Parental protein	Number of peptide fragments (percentage replicated compared with derivation dataset)		
		Derivation dataset (2005-2010)	Time-shifted dataset (2009-2013)	Synchronous dataset (2005-2010)
e01132	MGP	2	1 (50%)	2 (100%)
e05288	COL7A1	1	0 (0%)	0 (0%)
e02858	FRMD4A	1	0 (0%)	1 (100%)
e19279	PK1L2	1	1 (100%)	1 (100%)
e17095	UBAC2	1	0 (0%)	0 (0%)
e02105	COL4A2	1	0 (0%)	1 (100%)
e17751	COL14A1	1	1 (100%)	0 (0%)
e00579	PPP1R3A	1	0 (0%)	1 (100%)

Replication indicates directionally similarly associated with chronological age with a p value of <0.1 (given the prior probability established by the significance in the derivation dataset).

Table 5: Weights assigned by elastic net regression to the multidimensional UPP biomarker reflecting ageing

Peptide ID number	Amino-acid sequence	Parental protein identified from the amino-acid sequence		Weights β (95% confidence interval)
		Accession number	Name	
...	58.92735 (intercept)
e12949	LDGAKGDAGPAGpK GEpGSpGENGApG	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.009196 (-0.016636, -0.001756)
e16219	ppGAKGQEGAHGAP GAAGNPGAPGHVGA PGPSGpP	COL22A1 (Q8NFW1)	collagen type XXII alpha 1 chain	-0.005926 (-0.010442, -0.001410)
e11008	pPGEEGKRGRGDp GTVGPpGP	COL5A2 (P05997)	collagen type V alpha 2 chain	-0.004879 (-0.010490, 0.000732)
e16966	PpGESGREGApGAEG SpGRDGSPGAKGDR GETGP	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.003900 (-0.005744, -0.002056)
e15467	GRDGNpGNDGPPpGR DGQpGHKGERGYpG	COL1A2 (P08123)	collagen type I alpha 2 chain	-0.003467 (-0.006838, -0.000096)
e06650	EGSpGRDGSpGAKG DRG	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.003427 (-0.011087, 0.004233)
e18250	LALSGEKGDQGPPG DPGSPGSPGAPGPA GPPGYGpQGEP	COL4A3 (Q01955)	collagen type IV alpha 3 chain	-0.002831 (-0.005216, -0.000446)
e03183	KPWAAQDGPKPG	LRRC25 (Q8N386)	leucine rich repeat containing 25	-0.002656 (-0.008705, 0.003393)
e09435	GEKGPSGEAGTAGPp GTpGPQG	COL1A2 (P08123)	collagen type I alpha 2 chain	-0.002555 (-0.006981, 0.001871)
e10072	GLpGTGGPpGENGKp GEpGPKG	COL3A1 (P02461)	collagen type III alpha 1 chain	-0.002533 (-0.005587, 0.000521)
e04419	ApGDRGEpGPpGPAG	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.002351 (-0.003339, -0.001363)
e10881	NGEpGGKGERGAPG EKGEggppG	COL3A1 (P02461)	collagen type III alpha 1 chain	-0.002188 (-0.003979, -0.000397)
e10141	PpGEAGKPGEQGVp GDLGApGP	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.002027 (-0.004795, 0.000741)

Peptide ID number	Amino-acid sequence	Parental protein identified from the amino-acid sequence		Weights β (95% confidence interval)
		Accession number	Name	
e16540	GPAGAAGARGNDGQ PGPAGPPGVPgAG GpGFpGAPG	COL2A1 (P02458)	collagen type II alpha 1 chain	-0.001541 (-0.004379, 0.001297)
e10025	NSGEpGApGSKGDTG AKGEpGP	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.001431 (-0.003144, 0.000282)
e08755	EGSpGRDGSpGAKG DRGET	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.001387 (-0.003039, 0.000265)
e06709	GSpGSpGPDGKTGPp GPAG	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.001228 (-0.002277, -0.000179)
e12851	LDGAKGDAGPAGPK GEpGSpGENGApG	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.001142 (-0.005999, 0.003715)
e03016	ApGDRGEPGPpGP	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.001077 (-0.002990, 0.000836)
e02667	SpGPDGKTGPpPGP	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.001020 (-0.002335, 0.000295)
e06256	TGLSMDGGGSPKGD VDP	FXD2 (P54710)	FXD domain containing ion transport regulator 2	-0.000890 (-0.003897, 0.002117)
e17787	GARGLpGpPGSNGNP GPPGSPGKDGp PGPAGNTGAP	COL3A1 (P02461)	collagen type III alpha 1 chain	-0.000889 (-0.001681, -0.000097)
e15323	VQGPpGpAGEEGKR GARGEPTGLpGPp G	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.000788 (-0.002701, 0.001125)
e05259	DGQPGAKGEpGDAG AK	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.000745 (-0.002142, 0.000652)
e08188	GNDGApGKNGERGG pGGpGP	COL3A1 (P02461)	collagen type III alpha 1 chain	-0.000644 (-0.001751, 0.000463)
e10554	EGSpGRDGSpGAKG DRGETGPA	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.000592 (-0.000923, -0.000261)
e05560	GSpGSpGPDGKTGPP Gp	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.000469 (-0.001290, 0.000352)
e18041	NTGAPGSpGVSpKG DAGQpGEKGSpGAQ GppGAPGPLG	COL3A1 (P02461)	collagen type III alpha 1 chain	-0.000247 (-0.000405, -0.000089)

Peptide ID number	Amino-acid sequence	Parental protein identified from the amino-acid sequence		Weights β (95% confidence interval)
		Accession number	Name	
e15237	ERGEAGIpGVpGAKG EDGKDGSpGEpGAN G	COL3A1 (P02461)	collagen type III alpha 1 chain	-0.000221 (-0.000448, 0.000006)
e11073	NSGEpGApGSKGDTG AKGEPGpVG	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.000214 (-0.000441, 0.000013)
e01758	ApGDRGEpGpp	COL1A1 (P02452)	collagen type I alpha 1 chain	-0.000147 (-0.000574, 0.000280)
e19868	EEKAVADTRDQADGS RASVDSGSSEEQGG SSRALVSTLVPL	PIGR (P01833)	polymeric immunoglobulin receptor	0.000228 (-0.000274, 0.000730)
e00340	cDDYRLc	MGP (P08493)	matrix Gla protein	0.000256 (-0.004713, 0.005225)
e01132	cDDYRLcE	MGP (P08493)	matrix Gla protein	0.000692 (-0.000233, 0.001617)
e11452	ADGQpGAKGEpGDA GAKGDAGPPGp	COL1A1 (P02452)	collagen type I alpha 1 chain	0.000735 (0.000174, 0.001296)
e18377	FAEEKAVADTRDQAD GSRASVDSGSSEEQ GGSSRA	PIGR (P01833)	polymeric immunoglobulin receptor	0.000880 (-0.000684, 0.002444)
e04664	KSqAqSTDSSkDP	KRT12 (Q99456)	keratin 12	0.000892 (-0.000596, 0.002380)
e18492	DLADGVSGGEGKGG SDGGGSHRKEGEEA DAPGVIPGIVG	CD99 (P14209)	CD99 molecule (Xg blood group)	0.000930 (-0.000693, 0.002553)
e15070	YGGSSGGSSGGG YGGSSSGGHGGSS SGGYGGGS	KRT10 (P13645)	keratin 10	0.001110 (0.000308, 0.001912)
e00966	YQTNKAKH	CSTB (P04080)	cystatin B	0.001170 (-0.002740, 0.005080)
e10483	GPpGPpGKNGDDGE AGKpGRpG	COL1A1 (P02452)	collagen type I alpha 1 chain	0.001437 (-0.000153, 0.003027)
e17254	ENGKPGEpGpKGDA GApGApGGKGDAGAp GERGpPG	COL3A1 (P02461)	collagen type III alpha 1 chain	0.001536 (0.000062, 0.003010)
e08360	TGYTGKMcESSVNYc E	FAT4 (Q6V0I7)	protocadherin fat 4	0.001842 (0.001209, 0.002475)

Peptide ID number	Amino-acid sequence	Parental protein identified from the amino-acid sequence		Weights β (95% confidence interval)
		Accession number	Name	
e20087	ARGNDGARGSDGQp GpPGPpGTAGFpGSp GAKGEVpAGSpGSN GApG	COL3A1 (P02461)	collagen type III alpha 1 chain	0.001871 (0.000758, 0.002984)
e08916	DDGEAGKPGRpGER GpPGP	COL1A1 (P02452)	collagen type I alpha 1 chain	0.002693 (0.000829, 0.004557)
e09874	DGESGRpGRPGERG LpGPPG	COL3A1 (P02461)	collagen type III alpha 1 chain	0.002974 (0.000406, 0.005542)
e14837	ERGSpGPAGPKGSpG EAGRpGEAGLpGAKG	COL1A1 (P02452)	collagen type I alpha 1 chain	0.003067 (0.001293, 0.004841)
e11537	GKNGDDGEAGKPGR pGERGpGP	COL1A1 (P02452)	collagen type I alpha 1 chain	0.003790 (0.000489, 0.007091)
e05671	KpGEQGVpGDLGAPG p	COL1A1 (P02452)	collagen type I alpha 1 chain	0.003797 (0.001359, 0.006235)
e07799	GPpGEAGKpGEQGV PGDLG	COL1A1 (P02452)	collagen type I alpha 1 chain	0.003847 (0.001072, 0.006622)
e09697	PGpPGPHGPPGpmG PHGLpGP	COL25A1 (Q9BXS0)	collagen type XXV alpha 1 chain	0.005429 (0.003867, 0.006991)
e17505	AGRpGEVGPpGPpGP AGEKGSPPGADGPAG APGTpGPQG	COL1A1 (P02452)	collagen type I alpha 1 chain	0.005701 (-0.002982, 0.014384)
e09521	EGGKGAAGpPGPPG AAGTpGLQG	COL3A1 (P02461)	collagen type III alpha 1 chain	0.005761 (0.000716, 0.010806)
e10386	DAGAPGApGGKGDA GApGERGpPG	COL3A1 (P02461)	collagen type III alpha 1 chain	0.006532 (0.003161, 0.009903)

The derivation study included the baseline data of 559 FLEMENGHO participants. Accession numbers refer to the Uniprot database (<http://www.uniprot.org/uniprot>). Association sizes (β) between age and the abundance of the urinary peptides were derived by ELASTIC net regression, which was bootstrapped 1000 times. The regression slopes (β s) reflect the weight of each of the 54 peptide fragments in the construction of the multidimensional UPP marker. Of 210 peptides entering the data reduction analysis, 54 entered at least 60% of models and identified 17 proteins. The 95% confidence intervals of the β s were derived from the bootstrap distribution.

Table 6: Risk biomarkers in relation to age in 778 FLEMENGHO examined from 2005-2010 with adjustment for clustering within families

Risk biomarkers	Intrafamilial correlation		C-age		UPP-age		UPP-age-R	
	ICC	p value	r	p value	r	p value	r	p value
Vascular biomarkers								
Systolic pressure	0.14	0.00052	0.50	<0.0001	0.47	<0.0001	0.06	0.16
Diastolic pressure	0.31	<0.0001	0.24	<0.0001	0.23	<0.0001	0.04	0.39
VEGF	0.41	<0.0001	0.15	0.00020	0.16	0.00010	0.04	0.28
PAI-1	0.24	<0.0001	0.07	0.085	0.09	0.037	0.05	0.27
Renal biomarkers								
eGFR	0.19	<0.0001	-0.73	<0.0001	-0.66	<0.0001	-0.05	0.23
NGAL	0.25	<0.0001	0.17	<0.0001	0.19	<0.0001	0.09	0.036
Insulin resistance								
Plasma glucose	0.33	<0.0001	0.27	<0.0001	0.26	<0.0001	0.05	0.21
Serum insulin	0.14	0.00053	0.00	0.99	0.04	0.36	0.07	0.085
HOMA-IR	0.07	0.11	0.05	0.26	0.07	0.071	0.06	0.12
Lipid biomarkers								
Body mass index	0.28	<0.0001	0.34	<0.0001	0.34	<0.0001	0.10	0.021
Waist circumference	0.28	<0.0001	0.36	<0.0001	0.36	<0.0001	0.09	0.022
Body fat	-0.07	0.11	0.39	<0.0001	0.38	<0.0001	0.08	0.069
Leptin	0.20	<0.0001	0.07	0.11	0.11	0.011	0.09	0.027
Resistin	0.27	<0.0001	0.04	0.31	0.04	0.28	0.01	0.72
Inflammation biomarkers								
C-reactive protein	0.08	0.067	0.10	0.016	0.11	0.0097	0.04	0.36
TNF α	0.23	<0.0001	0.25	<0.0001	0.25	<0.0001	0.07	0.10
TNFR-1	0.24	<0.0001	0.36	<0.0001	0.34	<0.0001	0.06	0.15
Other markers								
Daily tobacco use in smokers	0.19	<0.0001	-0.06	0.18	-0.03	0.43	0.03	0.46
FSH in 395 women	0.00	0.98	0.66	<0.0001	0.55	<0.0001	0.01	0.87
Framingham risk score	0.11	0.0055	0.88	<0.0001	0.76	<0.0001	-0.01	0.78
dp-ucMGP	0.13	0.0010	0.41	<0.0001	0.41	<0.0001	0.10	0.019

Abbreviations: ICC, intraclass correlation coefficient (Donner A, Koval JJ. The estimation of intraclass correlation in the analysis of family data. *Biometrics*. 1980;36:19-25.); VEGF, vascular endothelial growth factor; PAI-1, plasminogen activator inhibitor-1; eGFR, glomerular filtration rate derived from serum creatinine; NGAL, neutrophil gelatinase-associated lipocalin; HOMA-IR, homeostatic model assessment insulin resistance; VCAM-1, vascular cell adhesion protein 1; VCAM-2, vascular cell adhesion protein 2; TNF α , tumour necrosis factor α ; TNFR-1, tumour necrosis factor receptor-1; FSH, follicle stimulating hormone; dp-ucMGP, desphospho-uncarboxylated matrix Gla protein. The distributions of VEGF, PAI-1, NGAL, HOMA-IR, body fat, visceral fat, leptin, resistin, C-reactive protein, TNF α , TNFR-1, daily use of smoking materials, FSH, Framingham risk score and dp-ucMGP were rank normalised. *r* is the Pearson correlation coefficient. C-age and UPP-age refer to chronological age and age as predicted by the UPP. UPP-age-R refers to the residual of the regression of UPP-age on C-age and reflects rapid (accelerated) ageing as predicted by the UPP-age, independent of C-age.

Table 7: Risk of adverse health outcomes in relation to age in 778 FLEMENGHO with adjustment for clustering within families

Endpoints	n/N (%)	C-age		UPP-age		UPP-age-R	
		HR (95%CI)	p value	HR (95%CI)	p value	HR (95%CI)	p value
Mortality							
Total	71/778 (9.1)	3.62 (2.87-4.57)	<0.0001	2.66 (2.22-3.19)	<0.0001	1.57 (1.20-2.05)	0.00096
Cardiovascular	24/778 (3.1)	7.27 (4.32-12.2)	<0.0001	3.15 (2.31-4.28)	<0.0001	1.81 (1.17-2.80)	0.0075
Non-cardiovascular	43/778 (5.5)	2.78 (2.10-3.67)	<0.0001	2.20 (1.78-2.72)	<0.0001	1.35 (0.94-1.92)	0.10
Fatal and non-fatal cancer	71/778 (9.1)	2.00 (1.67-2.40)	<0.0001	1.71 (1.47-1.98)	<0.0001	1.10 (0.81-1.49)	0.53
Musculoskeletal disorders							
Osteoarticular complications	162/778 (20.8)	1.03 (0.93-1.14)	0.61	1.07 (0.95-1.19)	0.26	1.15 (0.94-1.40)	0.18
Osteoporosis and fractures							
All	68/778 (8.7)	1.24 (1.06-1.46)	0.0083	1.34 (1.14-1.57)	0.00046	1.40 (1.06-1.85)	0.018
Women	48/395 (12.2)	1.55 (1.26-1.92)	0.0001	1.54 (1.28-1.85)	<0.0001	1.43 (1.03-1.99)	0.031
Men	20/383 (5.2)	0.82 (0.62-1.08)	0.16	0.86 (0.62-1.19)	0.36	1.19 (0.67-2.12)	0.55

n/N are the number of incident endpoint/number of participants at risk. Hazard ratios (HR), given with 95% confidence interval, express the relative risk per 10-year increment. The cause of death was not documented 4 participants. C-age and UPP-age refer to chronological age and age as predicted by the UPP. UPP-age-R refers to the residual of the regression of UPP-age on C-age and reflects rapid (accelerated) ageing as predicted by the UPP-age, independent of C-age.

Table 8: Pathway analysis

Pathway identifier	Pathway descriptor	q value	Proteins included
<i>Reactome</i>			
R-hsa-8948216	Collagen chain trimerisation	1.12E-35	COL25A1/ COL7A1/COL4A2/COL14A1/COL15A1/COL5A2/COL22A1/COL17A1/COL9A2/COL8A1/COL4A3/COL11A2/ COL2A1/COL1A2/COL4A1/COL19A1/ COL3A1/COL1A1
R-hsa-1650814	Collagen biosynthesis and modifying enzymes	4.96E-32	COL25A1/COL7A1/COL4A2/COL14A1/COL15A1/COL5A2/COL22A1/COL17A1/COL9A2/COL8A1/COL4A3/COL11A2/ COL2A1/COL1A2/COL4A1/COL19A1/COL3A1/COL1A1
R-hsa-1442490	Collagen degradation	3.97E-30	COL25A1/COL7A1/COL4A2/COL14A1/COL15A1/COL5A2/COL17A1/COL9A2/COL8A1/COL4A3/COL11A2/COL2A1/COL1 A2/COL4A1/COL19A1/COL3A1/COL1A1
R-hsa-1474290	Collagen formation	9.78E-30	COL25A1/COL7A1/COL4A2/COL14A1/COL15A1/COL5A2/COL22A1/COL17A1/COL9A2/COL8A1/COL4A3/COL11A2/COL2 A1/COL1A2/COL4A1/COL19A1/COL3A1/COL1A1
R-hsa-2022090	Assembly of collagen fibrils and multimeric structures	6.53E-26	COL7A1/COL4A2/COL14A1/COL15A1/COL5A2/COL17A1/COL9A2/COL8A1/COL4A3/COL11A2/COL2A1/COL1A2/COL4A 1/COL3A1/COL1A1
R-hsa-1474228	Degradation of the ECM	4.11E-24	COL25A1/COL7A1/COL4A2/COL14A1/COL15A1/COL5A2/COL17A1/COL9A2/COL8A1/COL4A3/COL11A2/COL2A1/COL1 A2/COL4A1/COL19A1/COL3A1/COL1A1
R-hsa-1474244	ECM organisation	7.74E-22	COL25A1/FGA/COL7A1/COL4A2/COL14A1/COL15A1/COL5A2/COL22A1/COL17A1/COL9A2/COL8A1/COL4A3/COL11A2/ COL2A1/COL1A2/COL4A1/COL19A1/COL3A1/COL1A1
R-hsa-216083	Integrin cell surface interactions	1.78E-17	FGA/COL7A1/COL4A2/COL5A2/COL9A2/COL8A1/COL4A3/COL2A1/COL1A2/COL4A1/COL3A1/COL1A1
R-hsa-3000171	Non-integrin membrane ECM interactions	2.34E-13	COL4A2/COL5A2/COL4A3/COL11A2/COL2A1/COL1A2/COL4A1/COL3A1/COL1A1
R-hsa-3000178	ECM protoglycans	2.31E-12	COL4A2/COL5A2/COL9A2/COL4A3/COL2A1/COL1A2/COL4A1/COL3A1/COL1A1
R-hsa-2214320	Anchoring fibril formation	8.13E-12	COL7A1/COL4A2/COL4A3/COL1A2/COL4A1/COL1A1
R-hsa-419037	NCAM1 interactions	8.63E-11	COL4A2/COL5A2/COL9A2/COL4A3/COL2A1/COL4A1/COL3A1
R-hsa-8874081	MET activates PTK2 signalling	7.93E-10	COL5A2/COL11A2/COL2A1/COL1A2/COL3A1/COL1A1
R-hsa-186797	Signalling by PDGF	7.99E-10	COL4A2/COL5A2/COL9A2/COL4A3/COL2A1/COL4A1/COL3A1
R-hsa-375165	NCAM signalling for neurite outgrowth	1.36E-09	COL4A2/COL5A2/COL9A2/COL4A3/COL2A1/COL4A1/COL3A1
R-hsa-2243919	Crosslinking of collagen fibrils	4.04E-09	COL4A2/COL4A3/COL1A2/COL4A1/COL1A1
R-hsa-8875878	MET promotes cell mobility	4.49E-09	COL5A2/COL11A2/COL2A1/COL1A2/COL3A1/COL1A1

Pathway identifier	Pathway descriptor	q value	Proteins included
R-hsa-3000480	Scavenging by class A receptors	4.87E-09	COL4A2/COL1A2/COL4A1/COL3A1/COL1A1
R-hsa-6806834	Signalling by MET	2.31E-07	COL5A2/COL11A2/COL2A1/COL1A2/COL3A1/COL1A1
R-hsa-2173782	Binding and uptake of ligands by scavenger receptors	3.06E-07	COL4A2/COL1A2/COL4A1/COL3A1/COL1A1
R-hsa-9006934	Signalling by receptor tyrosine kinases	1.08E-06	COL4A2/COL5A2/COL9A2/COL4A3/COL11A2/COL2A1/COL1A2/COL4A1/COL3A1/COL1A1
R-hsa-3000170	Syndecan interactions	2.42E-06	COL5A2/COL1A2/COL3A1/COL1A1
R-hsa-3000157	Laminin interactions	3.59E-06	COL7A1/COL4A2/COL4A3/COL4A1
R-hsa-198933	Immunoregulatory interactions between a lymphoid and non-lymphoid cell	3.88E-06	CD99/COL17A1/COL2A1/COL1A2/COL3A1/COL1A1
R-hsa-76009	Platelet aggregation (plug formation)	4.27E-04	FGA/COL1A2/COL1A1
R-hsa-76002	Platelet activation, signalling or aggregation	1.23E-02	FGA/PCDH7/COL1A2/COL1A1
R-hsa-202733	Cell surface interactions at the vascular wall	1.31E-02	CD99/COL1A2/COL1A1
KEGG			
hsa04974	Protein digestion and absorption	1.00E-29	COL25A1/COL7A1/COL4A2/COL14A1/FXYD2/COL15A1/COL5A2/COL22A1/COL17A1/COL9A2/COL8A1/COL4A3/COL11A2/COL2A1/COL1A2/COL4A1/COL19A1/COL3A1/COL1A1
hsa04512	ECM-receptor interaction	1.50E-07	COL4A2/COL9A2/COL4A3/COL2A1/COL1A2/COL4A1/COL1A1
hsa04933	AGE-RAGE signalling pathway in diabetic complications	5.89E-06	COL4A2/COL4A3/COL1A2/COL4A1/COL3A1/COL1A1
hsa05146	Amoebiasis	5.89E-06	COL4A2/COL4A3/COL1A2/COL4A1/COL3A1/COL1A1
hsa04510	Focal adhesion	1.56E-05	COL4A2/COL9A2/COL4A3/COL2A1/COL1A2/COL4A1/COL1A1
hsa04926	Relaxin signalling pathway	1.56E-05	COL4A2/COL4A3/COL1A2/COL4A1/COL3A1/COL1A1
hsa05165	Human papilloma virus infection	3.21E-04	COL4A2/COL9A2/COL4A3/COL2A1/COL1A2/COL4A1/COL1A1

Pathway identifier	Pathway descriptor	q value	Proteins included
hsa04151	PI3K-Akt signalling pathway	4.28E-04	COL4A2/COL9A2/COL4A3/COL2A1/COL1A2/COL4A1/COL1A1
hsa04611	Platelet activation	2.22E-03	FGA/COL1A2/COL3A1/COL1A1
hsa05222	Small cell lung cancer	9.72E-03	COL4A2/COL4A3/COL4A1

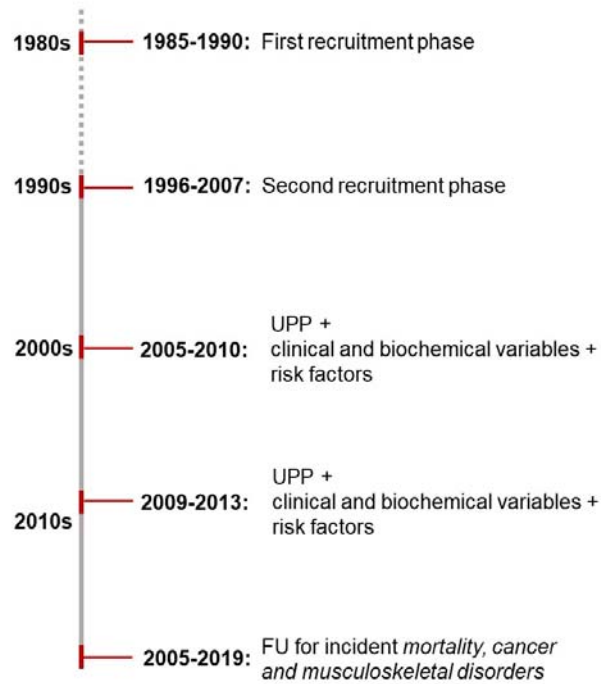
Abbreviations: ECM, extracellular matrix; NCAM1, neural adhesion molecule-1; MET, MET proto-oncogene, receptor tyrosine kinase; PTK2, protein tyrosine kinase 2; PDGF, platelet derived growth factor; AGE-RAGE, advanced glycation end product-receptor for advanced glycation end products; PI3K, phosphatidylinositol 3-kinase; Akt, protein kinase B. Pathway analysis was performed using the databases of Reactome (<https://reactome.org>) and the Kyoto Encyclopaedia of Genes and Genomes (<https://www.genome.jp/kegg>).

Table 9: Characteristics of FLEMENGHO participants analysed and not analysed

Characteristics	Analysed (n=778 of 778)	Not analysed (n=360 of 413)	p value
Number (%)			
Women	395 (50.8%)	186 (51.7%)	0.84
Smokers	176 (22.6%)	100 (27.8%)	0.063
Hypertension	278 (35.7%)	104 (28.9%)	0.026
Treated hypertension	146 (18.8%)	60 (16.7%)	0.41
Mean (SD) of characteristic			
Age, year	46.2 (16.2)	38.8 (19.2)	<0.0001
Body mass index, kg/m ²	25.7 (4.0)	25.7 (5.7)	0.84
Waist circumference, cm	83.1 (11.0)	82.9 (15.1)	0.77
Systolic pressure, mm Hg	128.1 (16.5)	124.5 (16.2)	0.00056
Diastolic pressure, mm Hg	78.9 (10.3)	75.9 (10.8)	<0.0001
Mean arterial pressure, mm Hg	95.3 (10.9)	92.1 (11.5)	<0.0001
Serum creatinine, µmol/l	86.8 (14.7)	87.7 (17.7)	0.35
eGFR, ml/min/1.73 m ²	85.1 (19.2)	88.8 (23.3)	0.0044
Plasma glucose, mmol/l	5.04 (0.98)	5.14 (1.55)	0.15
Total serum cholesterol, mmol/l	5.29 (1.09)	5.25 (1.14)	0.48
HDL serum cholesterol, mmol/l	1.40 (0.39)	1.37 (0.37)	0.23
Total-to-HDL cholesterol ratio	4.07 (1.42)	4.12 (1.50)	0.64
Median of characteristic (IQR)			
γ-glutamyltransferase, U/l	20.6 (13.0-30.0)	18.1 (12.0-25.0)	0.0016

Analysed refers to 778 FLEMENGHO participants examined in 2005-2010 and making up the derivation and synchronous internal validation datasets. Not analysed refers to: (i) participants ineligible for follow-up, because they had passed away (n=26), were incapacitated (n=27), or had moved out of the area (n=100); (ii) because they declined renewed participation (n=227); or because they were excluded from analysis, because of missing UPP data (n=24) or because of missing (n=9) or unreliably recorded (n=17) covariables. The characteristics listed were extracted from a foregoing examination (1996-2005). The earlier data were available for all analysed participants but only for 360 of 413 participants (87.2%) not analysed. Hypertension was a blood pressure of ≥ 140 mm Hg systolic or ≥ 90 mm Hg diastolic or use of antihypertensive drugs. The glomerular filtration rate was derived from serum creatinine, using the Chronic Kidney Disease Epidemiology Collaboration equation. To convert eGFR from ml/min to ml/s, multiply by 0.0167. p values refer to the difference between participants analysed and not analysed.

FLEMENGHO TIMELINES



Timelines

Abbreviations: FLEMENGHO, Flemish Study of Environment, Genes and Health Outcomes; UPP, urinary proteomic profiling; FU, follow-up.

Analysis strategy

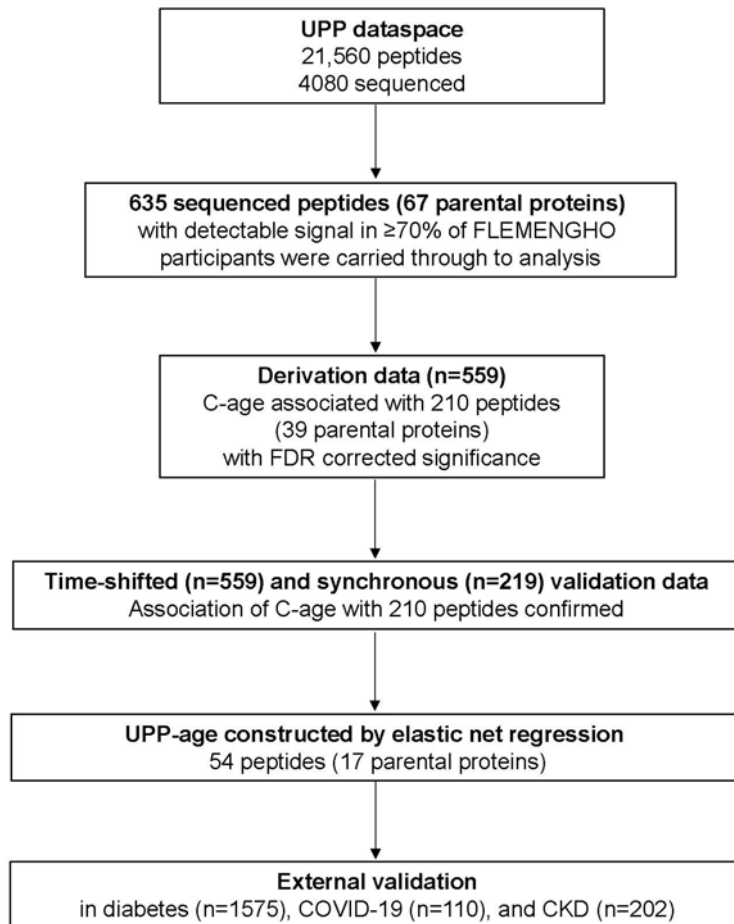


Figure 2:

Analysis strategy

Abbreviations: FLEMENGHO, Flemish Study on Environment, Genes and Health Outcomes; C-age, chronological age; FDR, false discovery rate; CKD, chronic kidney disease.

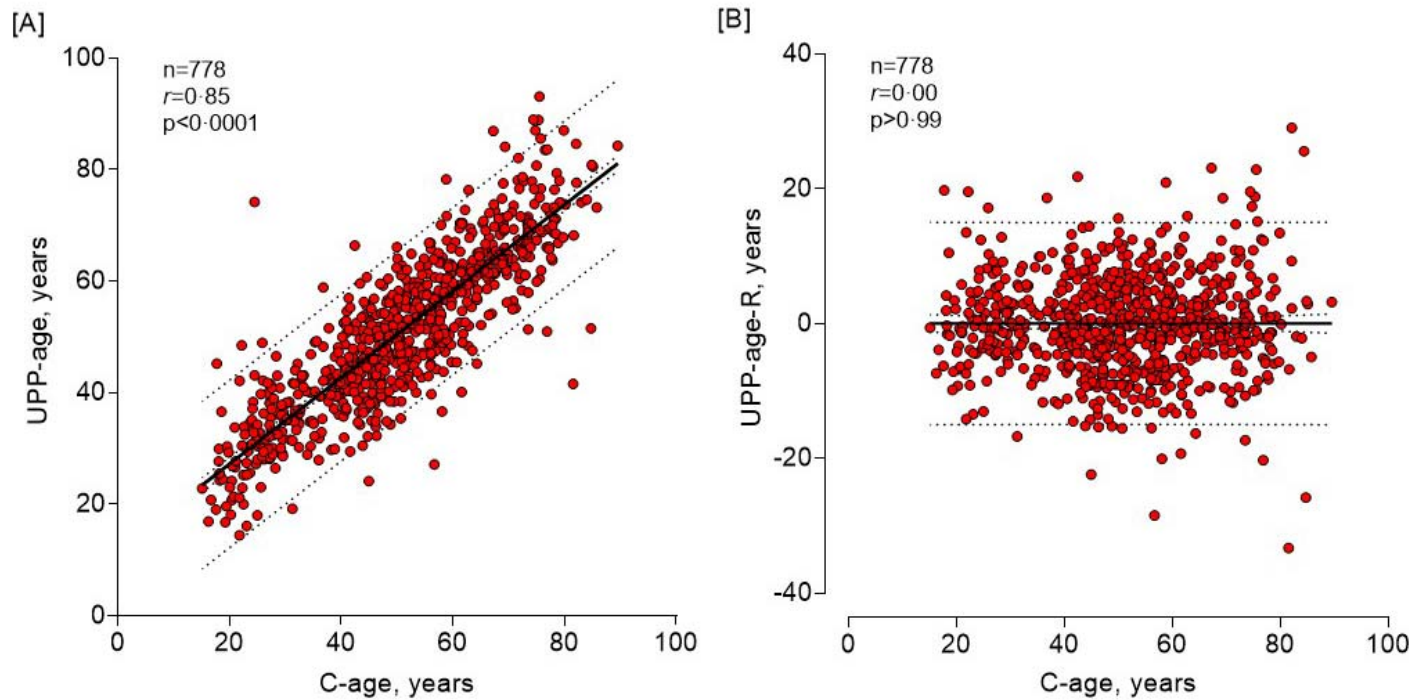


Figure 3:

Residual plot of UPP-age on C-age

UPP-age is associated with C-age (panel A). The residual plot of this association (panel B) shows that by removing the variability in UPP-age explained by C-age (UPP-age-R obtained by regressing UPP-age on C-age) removes the association, whereby the correlation coefficient drops from 0.85 ($p<0.0001$) to 0.00 ($p>0.99$). UPP-age-R therefore reflects the part of ageing that is not explained by chronological age (C-age).

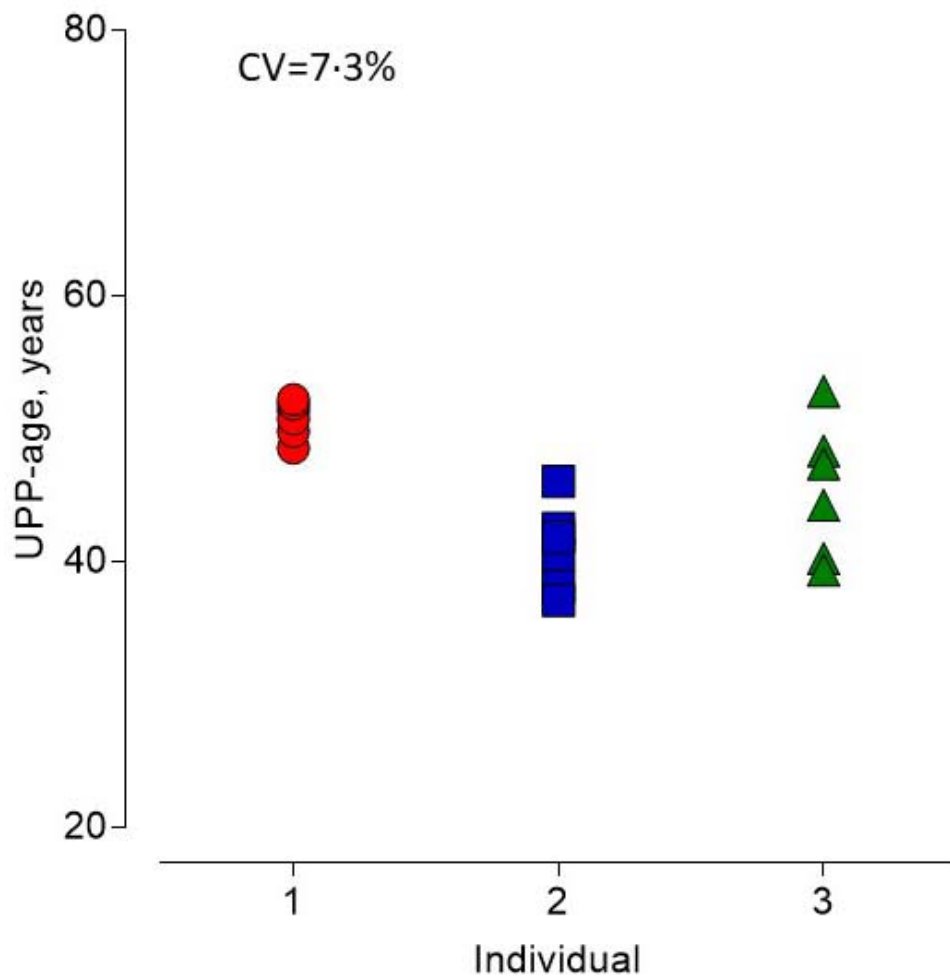


Figure 4:

Short-term reproducibility of the UPP ageing biomarker

The coefficient of variation of UPP-age was determined in three healthy individuals not included in the population or patient cohorts, who each underwent UPP profiling on mid-morning urine samples collected on six different days within a 1-month time span. The within individual coefficients of variation ranged from 2.7 to 11.3% and averaged (SE) 7.3 (2.5)%, indicating satisfactory short-term reproducibility.

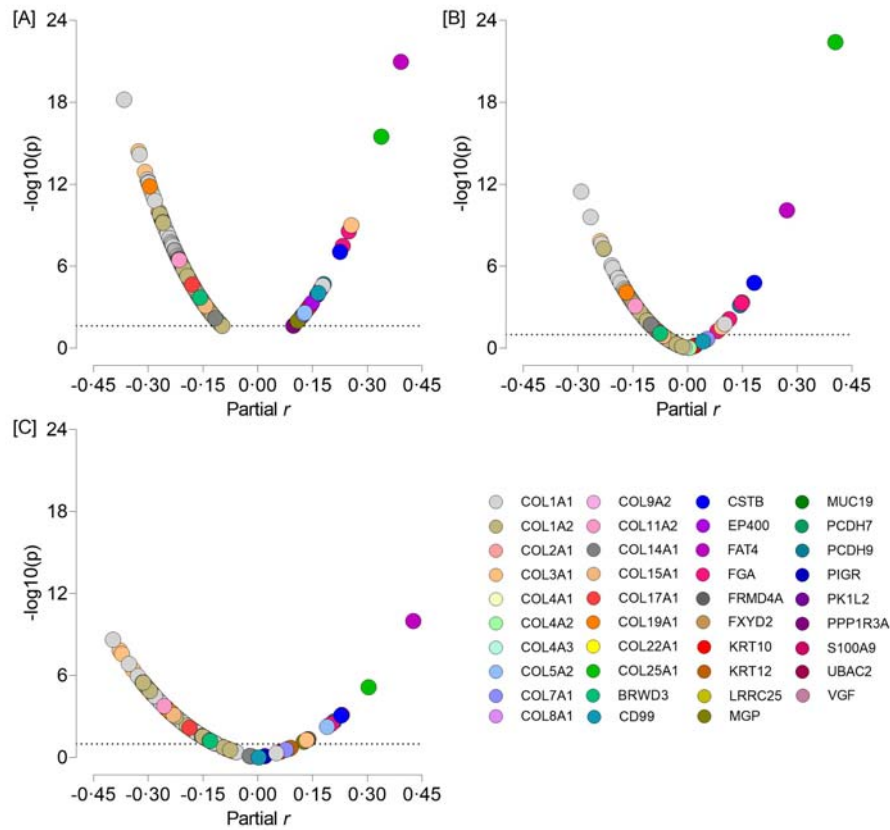


Figure 5:

Volcano plots for FLEMENGHO derivation (A) and time-shifted (B) and synchronous (C) internal validation datasets

Dotted lines indicate the level of significance required in the different stages of the analysis. The partial correlation coefficients plotted along the horizontal axis were derived from the regression model in table 2. The proteins identified from the peptide fragments were colour-coded as indicated in the legend.