# Supplemental Material

# Gut microbiota-generated phenylacetylglutamine and heart failure

Kymberleigh A. Romano PhD[1], Ina Nemet PhD[1], Prasenjit Prasad Saha PhD[1], Arash Haghikia MD[2,3,4], Xinmin S. Li PhD[1], Maradumane L. Mohan PhD[1], Beth Lovano BS[1], Laurie Castel BS[1], Marco Witkowski MD[1], Jennifer A. Buffa MS[1], Yu Sun PhD[1], Lin Li PhD[1], Christopher M. Menge BS[1], Ilja Demuth PhD[5,6], Maximilian König MD[5], Elisabeth Steinhagen-Thiessen MD[5], Joseph A. DiDonato PhD[1], Arjun Deb MD[7], Fredrik Bäckhed PhD[8], W.H. Wilson Tang MD[1,9], Sathyamangla Venkata Naga Prasad PhD[1], Ulf Landmesser MD[2,3,4], David R. Van Wagoner PhD[1], Stanley L. Hazen MD PhD[1,9]

[1] Department of Cardiovascular & Metabolic Sciences, Lerner Research Institute, Cleveland Clinic, Cleveland, OH, USA

[2] Department of Cardiology, Charité-Universitätsmedizin Berlin, Campus Benjamin Franklin, Berlin, Germany

[3] German Center for Cardiovascular Research (DZHK), Partner Site Berlin, Berlin, Germany

[4] Berlin Institute of Health (BIH), Berlin, Germany

[5] Department of Endocrinology and Metabolism, Charité-Universitätsmedizin Berlin, Charitéplatz, Berlin, Germany

[6] Berlin Institute of Health Center for Regenerative Therapies, Berlin, Germany

[7] Division of Cardiology and Department of Medicine, David Geffen School of Medicine, UCLA, Los Angeles, California, USA.

[8] Department of Molecular and Clinical Medicine, Wallenberg Laboratory, Institute of Medicine, University of Gothenburg, Gothenburg, Sweden.

[9] Heart, Vascular and Thoracic Institute, Cleveland Clinic, Cleveland, OH, USA

Lead contact: Stanley Hazen, 9500 Euclid Ave, mail code NC-10, Cleveland Clinic, Cleveland, OH, USA 44195

Tel: +1 216 445 9763. Fax: +1 216 444 9404.

Email: hazens@ccf.org

**Supplemental Table 1:** Baseline characteristics of heart failure (HF) patients in the US and European cohorts

| Characteristics | HF patients in the US Cohort (n=712) | HF patients in the European Cohort (n=553) |
|---|---|---|
| Age, mean ±SD, years | 66.1±10.5 | 76.1±9.5 |
| Male sex, n (%) | 421 (59.1) | 412 (74.3) |
| Current smoking, n (%) | 81 (11.4) | 69( 12.5) |
| Systolic blood pressure, mm Hg | 130.0 (116.0-145.0 ) | N/A |
| Diastolic blood pressure, mm Hg | 71.0 (63.0-80.0) | N/A |
| BMI, kg/cm$^2$ | 28.1 (25.1-32.9) | N/A |
| Hypertension, n (%) | 537 (76.8) | 460 (83.2) |
| Diabetes, n (%) | 287 (40.3) | 174 (31.5) |
| CAD, n (%) | 585 (82.2) | 460 (83.6) |
| HDL, mg/dL | 31.7 (26.0-39.7) | 47.0 (39.0-58.0) |
| LDL, mg/dL | 91.0 (72.8-112.0) | 87.0 (66.0-114.0) |
| TG, mg/dL | 118.5 (87.0-165.0) | 117.0 (89.0-167.0) |
| hs-CRP, mg/L | 3.90 (1.57-9.01) | 2.1 (0.8-5.6) |
| LVEF, % | 40.0 (25.0-55.0) | 56.0 (48.0-63.0) |
| NT-proBNP, pg/mL | 895.0 (348.8-2040.2) | 501.0 (173.0-1421.0) |
| eGFR, mL/min/1.73 m² | 79.8 (61.5-94.3) | 69.0 (56.8-85.6) |

The cohorts are comprised of sequential stable subjects without evidence of acute coronary syndrome (cardiac troponin I< 0.03 ng/mL) who underwent elective diagnostic coronary angiography (cardiac catheterization or coronary computed tomography) for evaluation of coronary artery disease (CAD).

Continuous data are presented as median (interquartile range or 25$^{th}$ percentile - 75$^{th}$ percentile), categorical variables are presented as %; N/A = not available; BMI = body mass index; HDL= high-density lipoprotein; LDL= low-density lipoprotein; TG = triglyceride; hs-CRP = high-sensitivity C-reactive protein; LVEF = left ventricular ejection fraction; NT-proBNP = N-terminal pro B-type natriuretic peptide; eGFR = estimated glomerular filtration rate.

The Wilcoxon–rank sum test or Welch two sample t-test for continuous variables and the $\chi 2$ test for categorical variables were used to determine significant difference between groups.

Estimated glomerular filtration rate (eGFR) (mL/min per 1.73 m$^2$) is calculated on the basis of the Chronic Kidney Disease Epidemiology Collaboration 2021 CKD-EPI Creatinine equation[15].

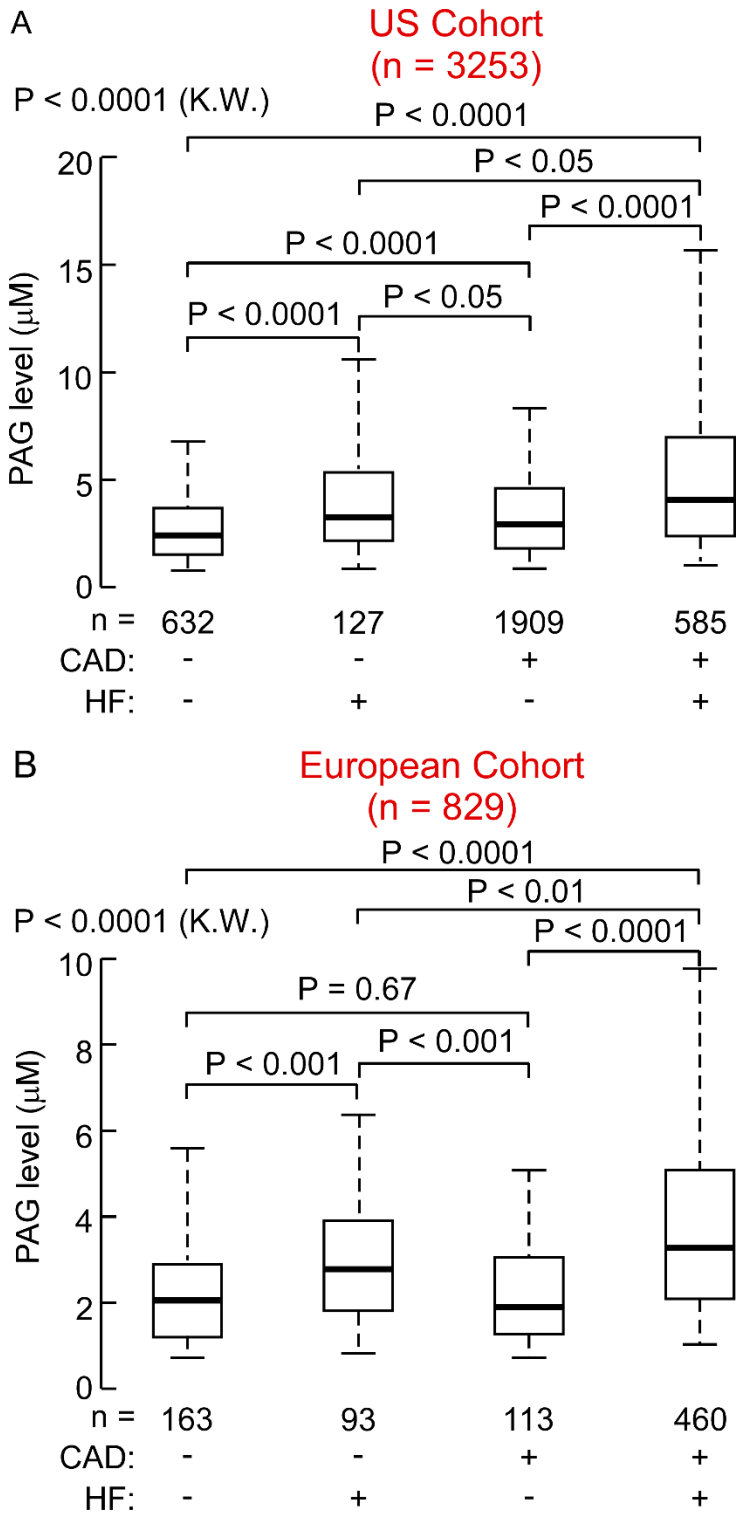



**Fig. S1. The association of PAGln with heart failure.** Box whisker plot of circulating PAGln levels in (**A**) US Cohort subjects or (**B**) European cohort subjects stratified by

coronary artery disease (CAD) and heart failure (HF) status. Data are represented as boxplots: middle line is the median, the lower and upper hinges are the first and third quartiles, the whiskers represent $10^{th}$ and $90^{th}$ percentile; P values were calculated using Kruskal Walls (K.W.) with Dunn's test.

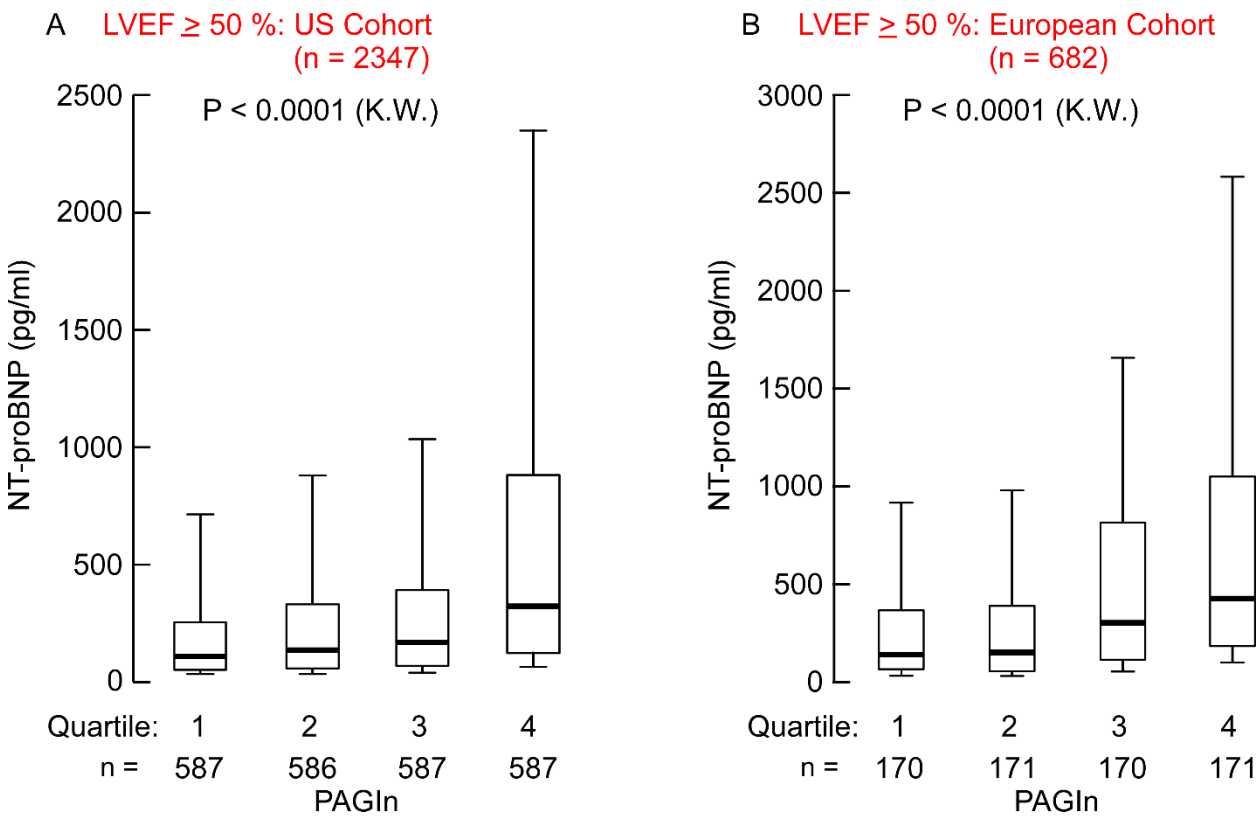



**Fig. S2. PAGln is associated with NT-proBNP in individuals with preserved left ventricular systolic function.**

Circulating NT-proBNP levels (pg/ml) among those with preserved left ventricle systolic ejection fraction (LVEF ≥ 50%) within (**A**) US Cohort; and (**B**) European Cohort stratified by PAGln quartiles. Data are represented as boxplots: middle line is the median, the lower and upper boundaries to the boxes represent 25$^{th}$ and 75$^{th}$ percentiles, and the whiskers represent 10$^{th}$ and 90$^{th}$ percentile; P values were calculated using Kruskal-Wallis (K.W.) test.

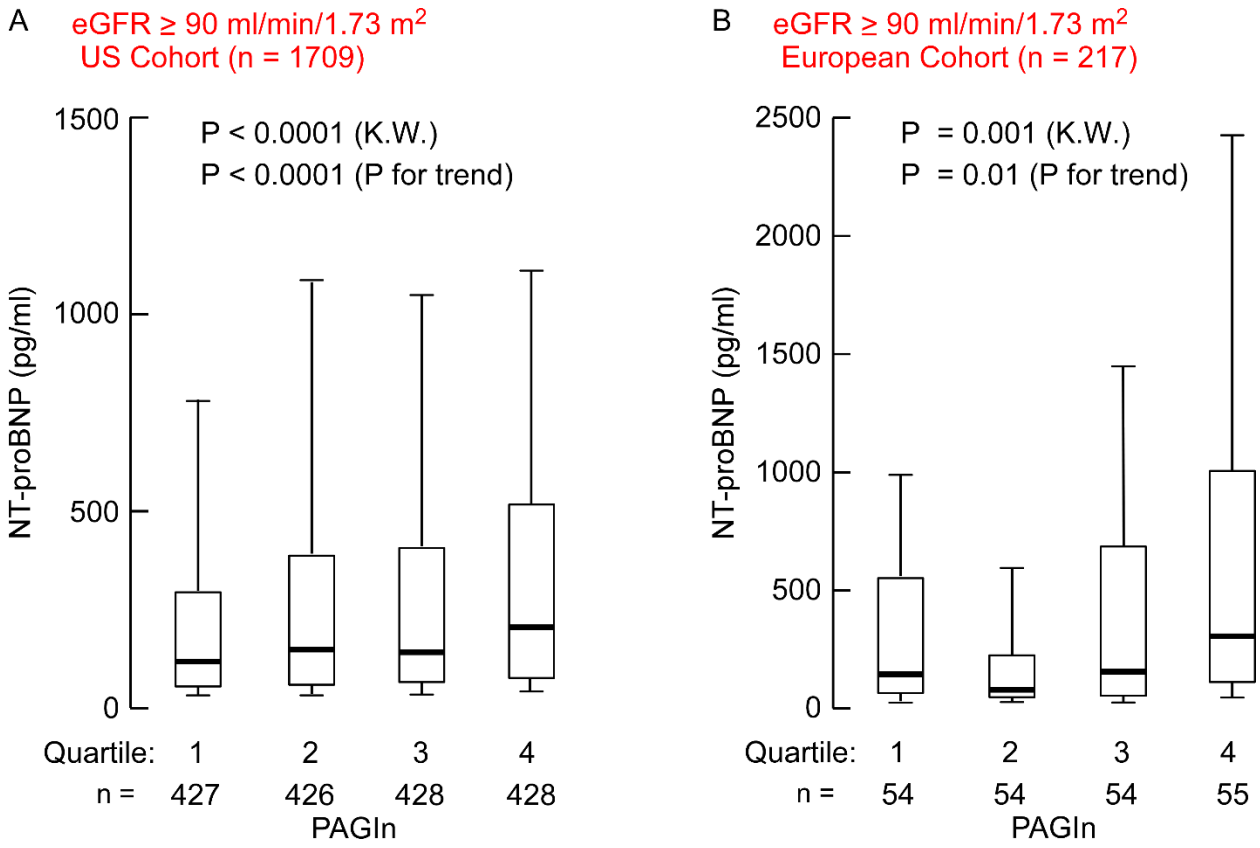
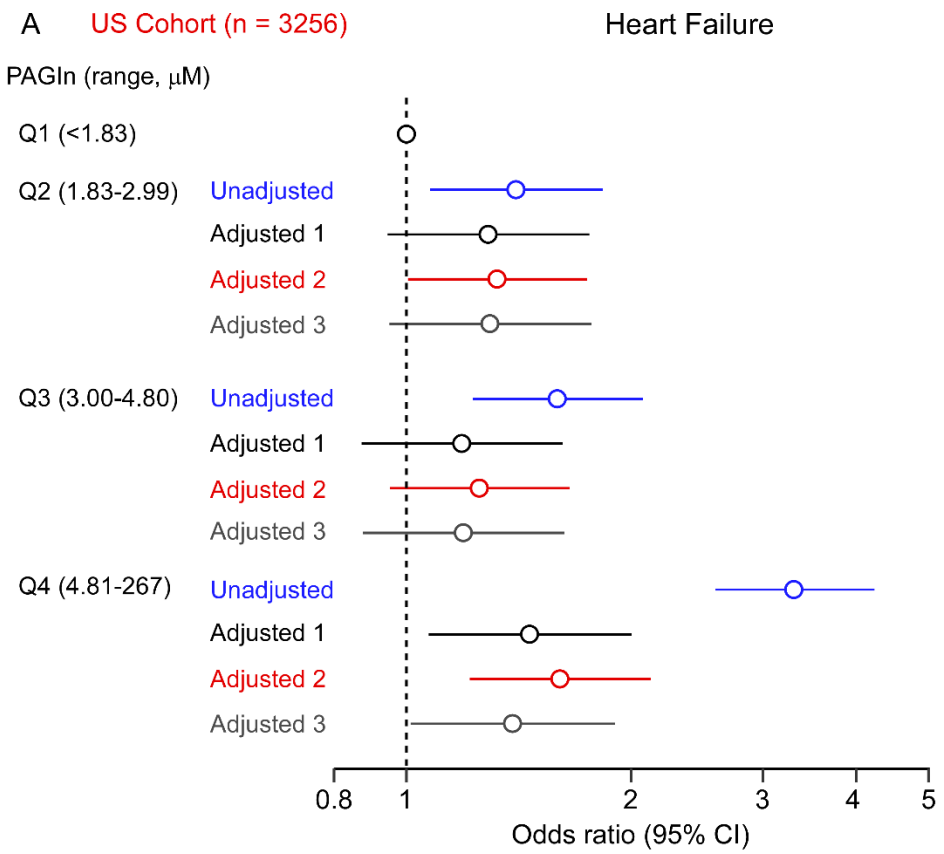



**Fig. S3. PAGln is associated with NTpro-BNP in individuals with preserved kidney function.**

Circulating NTpro-BNP levels (pg/ml) in (**A**) US Cohort or (**B**) European Cohort stratified by PAGln quartiles amongst those with normal renal function (eGFR≥90 mL/min/1.73 $m^2$). Data are represented as boxplots: middle line is the median, the lower and upper boundaries to the boxes represent 25$^{th}$ and 75$^{th}$ percentiles, and the whiskers represent 10$^{th}$ and 90$^{th}$ percentile; P values were calculated using Kruskal-Wallis (K.W.) test, and the Jonckheere –Terpstra test of trend.


A
US Cohort (n = 3256)
Heart Failure
PAGln (range, μM)
Q1 (<1.83)
Q2 (1.83-2.99)
Unadjusted
Adjusted 1
Adjusted 2
Adjusted 3
Q3 (3.00-4.80)
Unadjusted
Adjusted 1
Adjusted 2
Adjusted 3
Q4 (4.81-267)
Unadjusted
Adjusted 1
Adjusted 2
Adjusted 3
0.8
1
2
3
4
5
Odds ratio (95% CI)

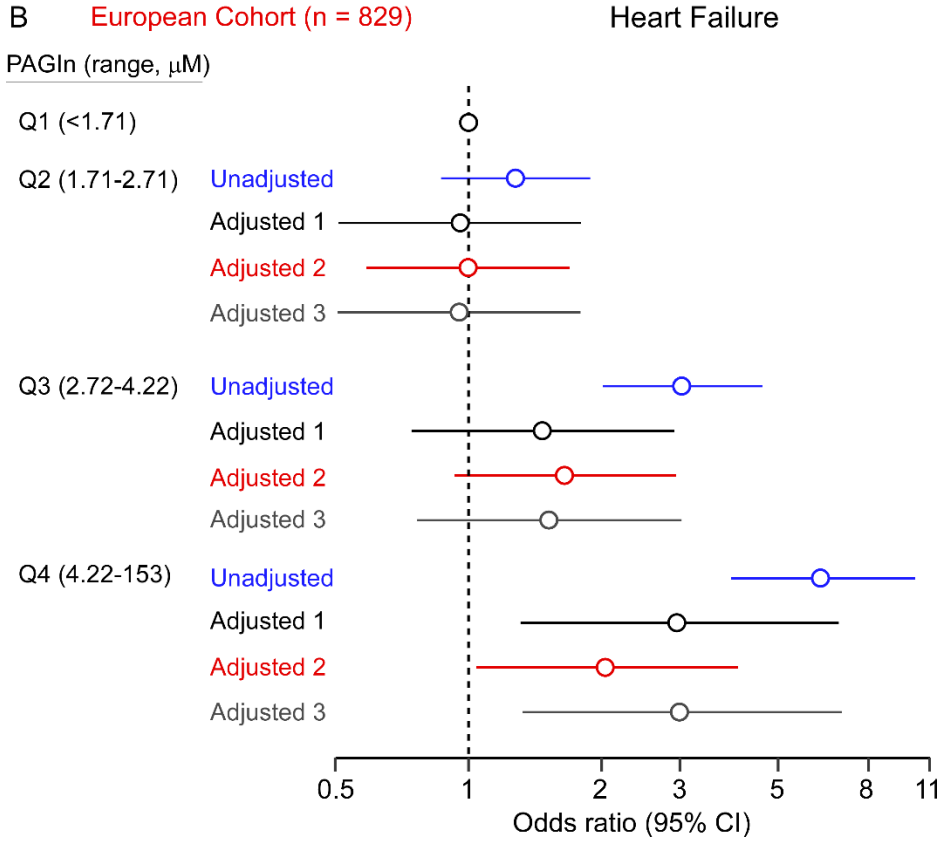

B
European Cohort (n = 829)
Heart Failure
PAGln (range, μM)
Q1 (<1.71)
Q2 (1.71-2.71)
Unadjusted
Adjusted 1
Adjusted 2
Adjusted 3
Q3 (2.72-4.22)
Unadjusted
Adjusted 1
Adjusted 2
Adjusted 3
Q4 (4.22-153)
Unadjusted
Adjusted 1
Adjusted 2
Adjusted 3
0.5
1
2
3
5
8
11
Odds ratio (95% CI)

**Figure S4. The association of PAGln with heart failure.**

**(A/B)** Risk of HF among all test subjects according to PAGln quartile levels using a multivariable logistic regression models. Unadjusted odd ratio (blue), Adjusted Model 1 (age, sex, smoking status, SBP, LDL, HDL, TG, hs-CRP, diabetes and obesity (BMI≥30 kg/$m^2$), indices of renal function (eGFR≥60, or <60 mL/min per 1.73 $m^2$) and LVEF, black); Adjusted Model 2 (age, sex, smoking status, SBP, LDL, HDL, TG, hs-CRP, diabetes and obesity (BMI≥30 kg/$m^2$), indices of renal function (eGFR≥60, or <60 mL/min per 1.73 $m^2$) and NT-proBNP, red); and Adjusted Model 3 (Model 2 + LVEF) . Symbols represent odds ratios, and the 5%–95% confidence intervals (CI) are indicated by the line length.