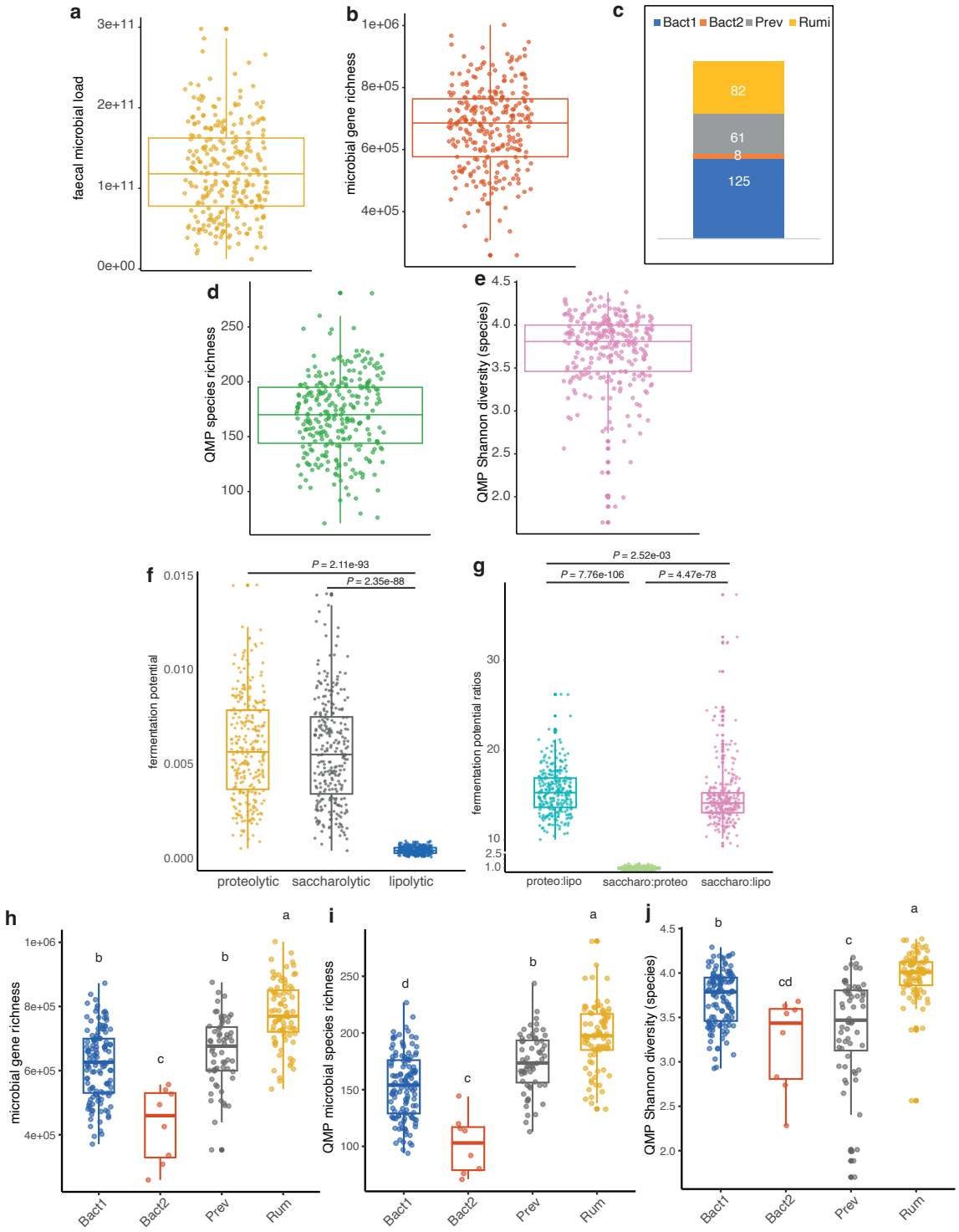
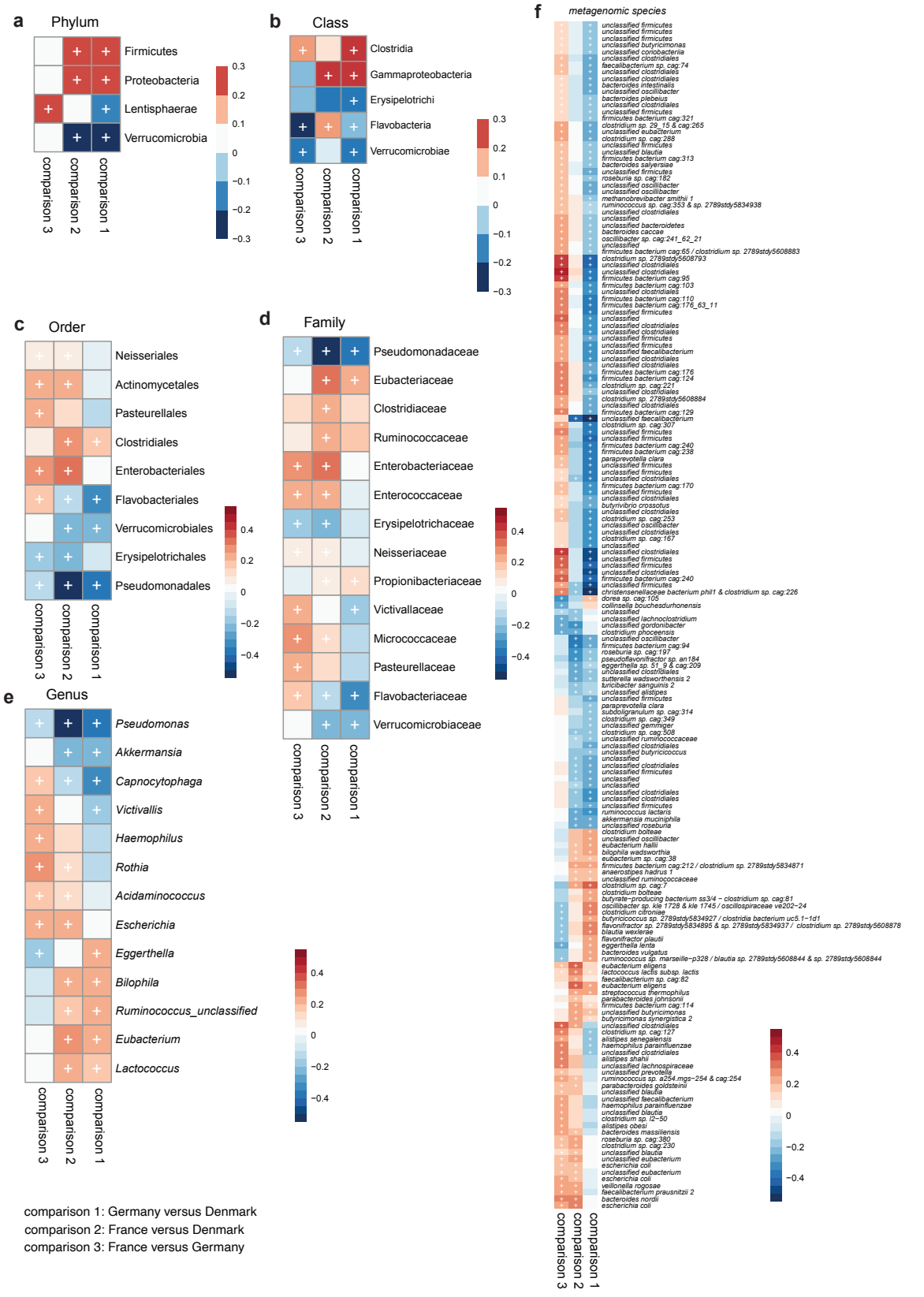


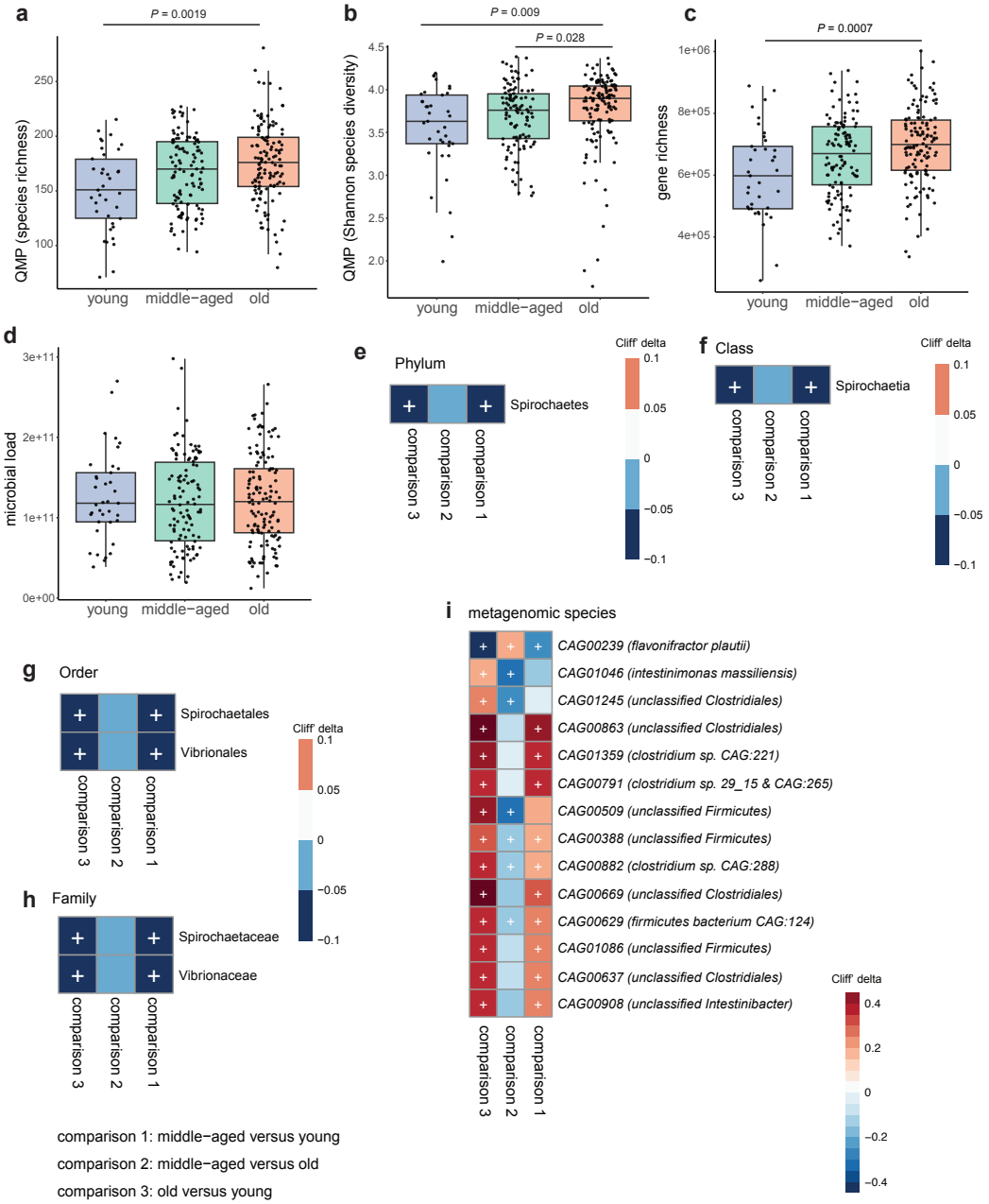
Supplementary Figure. 1. Metagenomic features of metabolically healthy individuals (n = 275). Distribution of **a**, faecal microbial load, **b**, microbial gene richness, **c**, enterotypes, **d,** microbial species richness and **e,** Shannon species diversity, **f**, saccharolytic-, proteolytic- and lipolytic- fermentation potential, and **g**, their respective ratios in the MetaCardis healthy individuals. *P*-values were derived from Kruskal-Wallis with Dunn-Bonferroni *post hoc* tests. Next, we tested the distribution of these variables based on enterotypes and only those exhibiting significant differences are shown here. Distribution of **h**, microbial gene richness **i,** Shannon species

richness, and **j,** microbial species diversity based on enterotypes in the clinically healthy individuals (Kruskal-Wallis with Dunn-Bonferroni *post hoc* tests; different alphabets represent $P \leq 0.05$). Bact1, *Bacteroides 1* enterotype; Bact2, *Bacteroides 2* enterotype; Prev, *Prevotella* enterotype; Rumi, *Ruminococcus* enterotype; QMP, quantitative microbiome profiling.



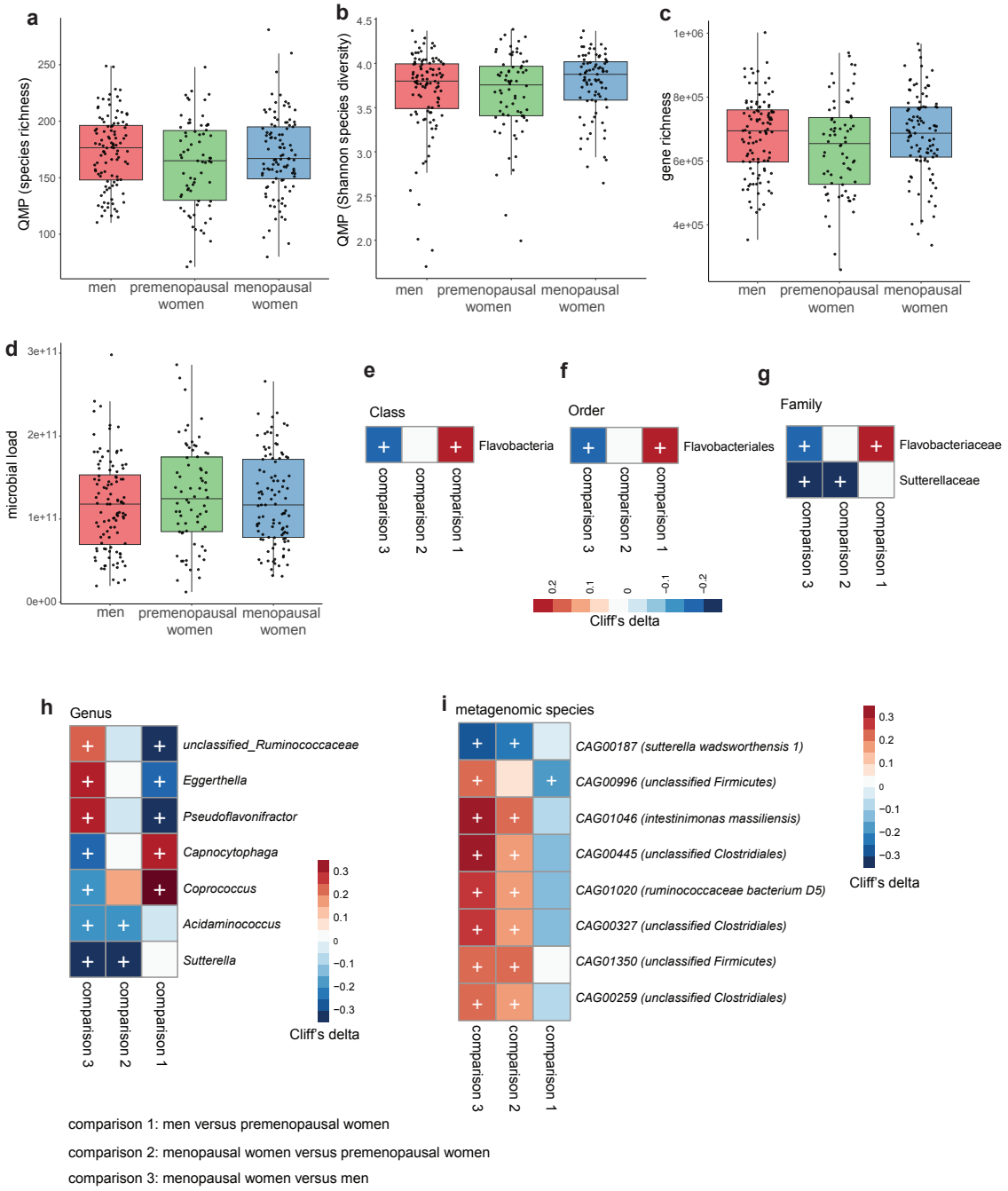
Supplementary Figure. 2. Impact of country of participant recruitment on the gut microbial composition (n = 275). Heatmaps showing the differences in the gut microbiome composition at the level of a) Phylum b) Class, c) Order, d) Family, e) Genus and f) Species

among participants recruited from France, Germany and Denmark within the healthy group (n = 275). Effect sizes were calculated using Cliff's Delta. *P*-values were derived from Kruskal-Wallis adjusted for multiple comparisons using the BH criteria. Only variables exhibiting FDR_{KW} < 0.1 were subjected to Dunn's pairwise comparisons adjusted for multiple comparisons using the BH criteria. + represents significance as FDR < 0.1 for each pairwise comparison.



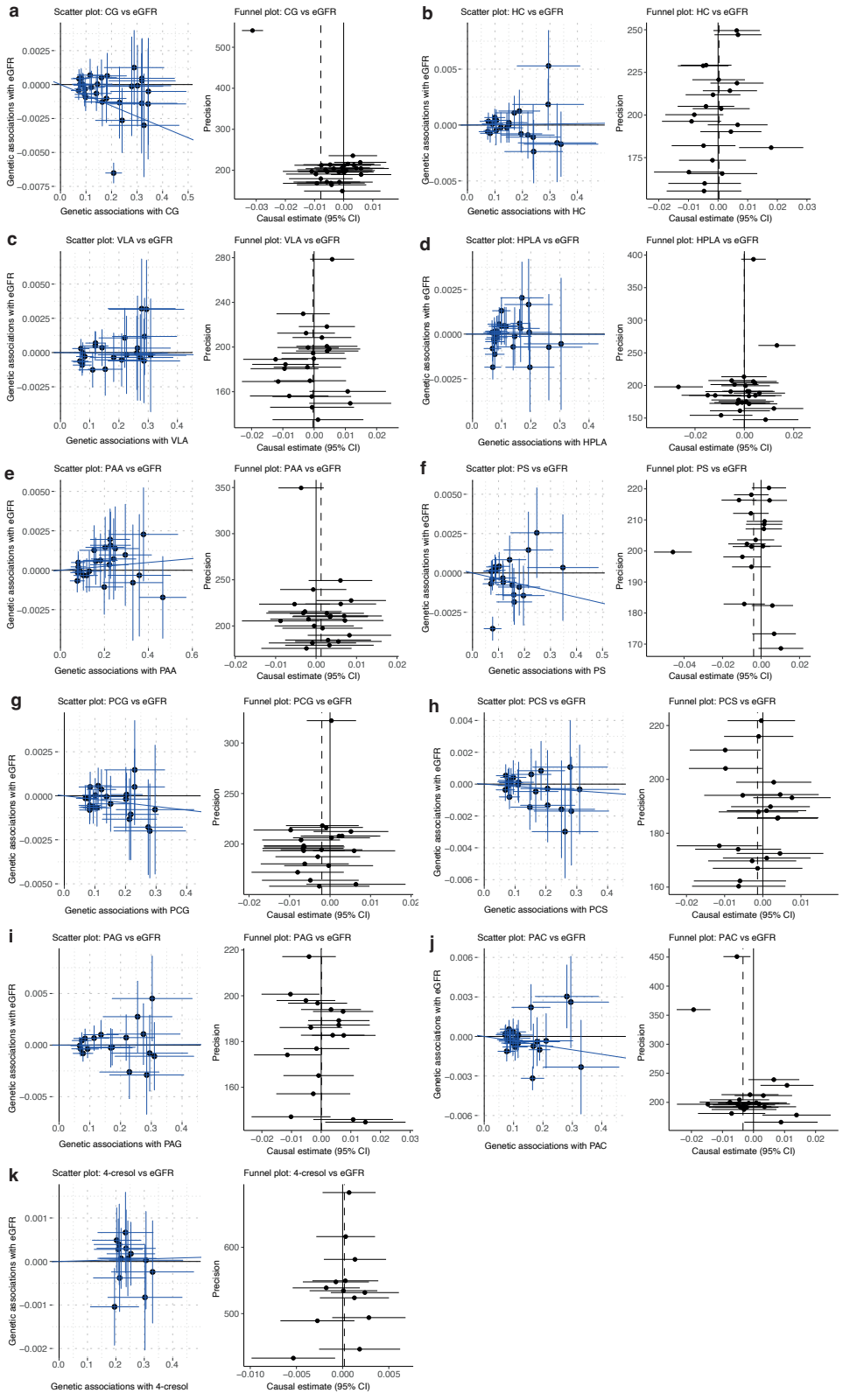
Supplementary Figure. 3. Impact of participant age on the gut microbial composition (n = 275). Box plots showing the differences in the **a,** microbial species richness, **b,** Shannon's species diversity, **c,** microbial gene richness and **d,** fecal microbial load in the healthy group participants when categorized as young (i.e., 20-40 years of age), middle-aged (i.e., 41-59 years of age) and old (i.e., 60-76 years of age). The body of the boxplot represents the first and third quartiles of the distribution with the median line, and the whiskers extend from the quartiles to the last data point within 1.5× interquartile range with outliers beyond. *P*-values were derived from Kruskal-Wallis followed by Dunn's *posthoc* tests. Heatmaps showing the impact of age of participants on the gut microbiome composition at the level of **e,** Phylum, **f,** Class, **g,** Order, **h,** Family and **i,** Species among the healthy individuals (n = 275). Effect sizes

were calculated using Cliff's Delta. *P*-values were derived from Kruskal-Wallis adjusted for multiple comparisons using the BH criteria. Only variables exhibiting $FDR_{KW} < 0.1$ were subjected to Dunn's pairwise comparisons, which were also adjusted for multiple comparisons using the BH criteria (n = 275). + represents significance as FDR < 0.1 for each pairwise comparison.



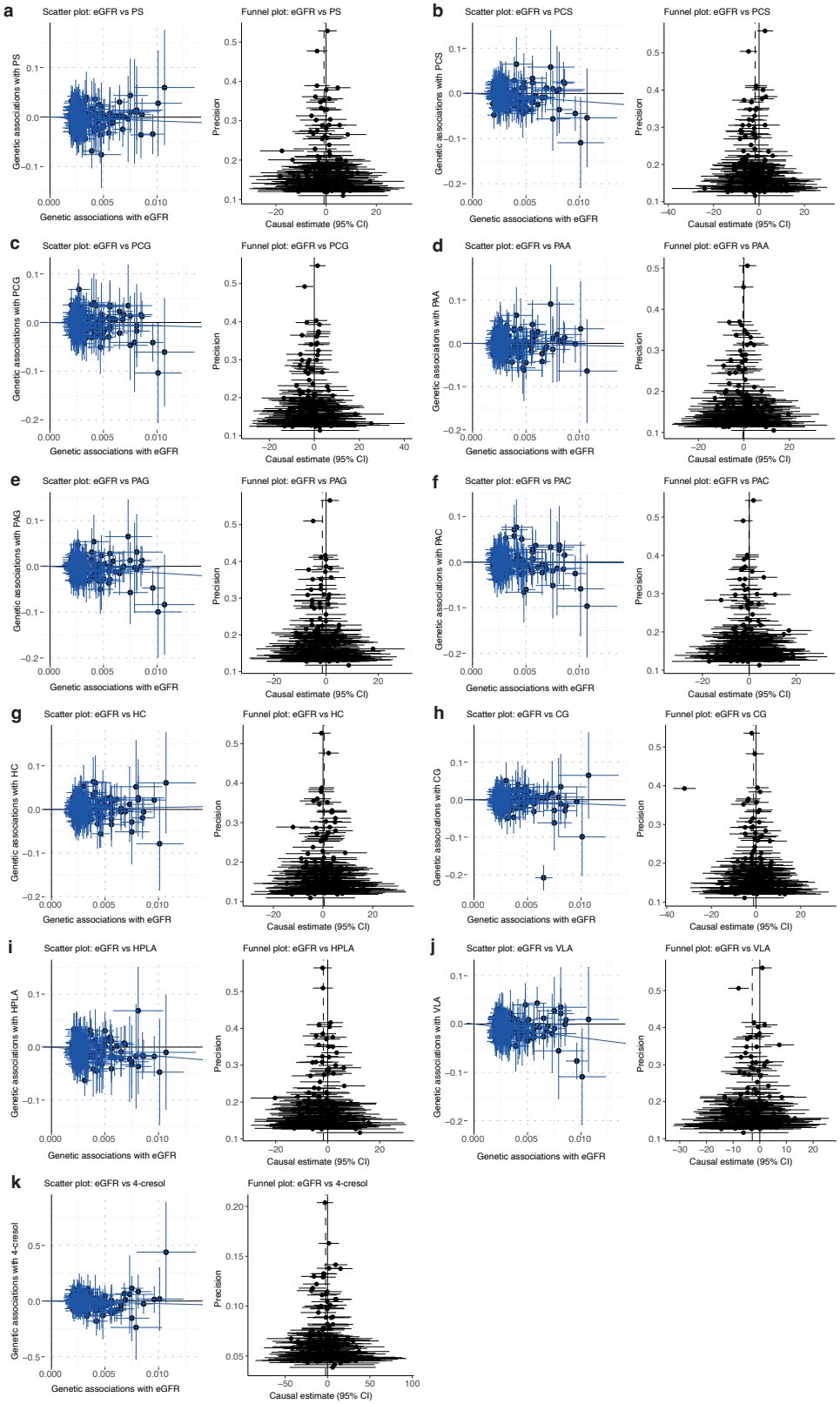
Supplementary Figure. 4. Impact of participant sex and menopausal status on the gut microbial composition (n = 275). Box plots showing the differences in the **a,** microbial species richness, **b,** Shannon's species diversity, **c,** microbial gene richness and **d,** fecal microbial load among men, premenopausal women and post-menopausal women within the healthy group (n = 275).The body of the boxplot represents the first and third quartiles of the distribution with the median line, and the whiskers extend from the quartiles to the last data point within 1.5× interquartile range with outliers beyond. *P*-values were derived from Kruskal-

Wallis followed by Dunn's *posthoc* tests. Heatmaps showing the differences in the gut microbiome composition at the level of **e,** Class, **f,** Order, **g,** Family **h,** Genus and **i,** Species among men, premenopausal women and post-menopausal women within the metabolically healthy group (n = 275). Effect sizes were calculated using Cliff's Delta. *P*-values were derived from Kruskal-Wallis adjusted for multiple comparisons using the BH criteria. Only variables exhibiting $FDR_{KW} < 0.1$ were subjected to Dunn's pairwise comparisons adjusted for multiple comparisons using the BH criteria (n = 275). + represents significance as FDR < 0.1 for each pairwise comparison.



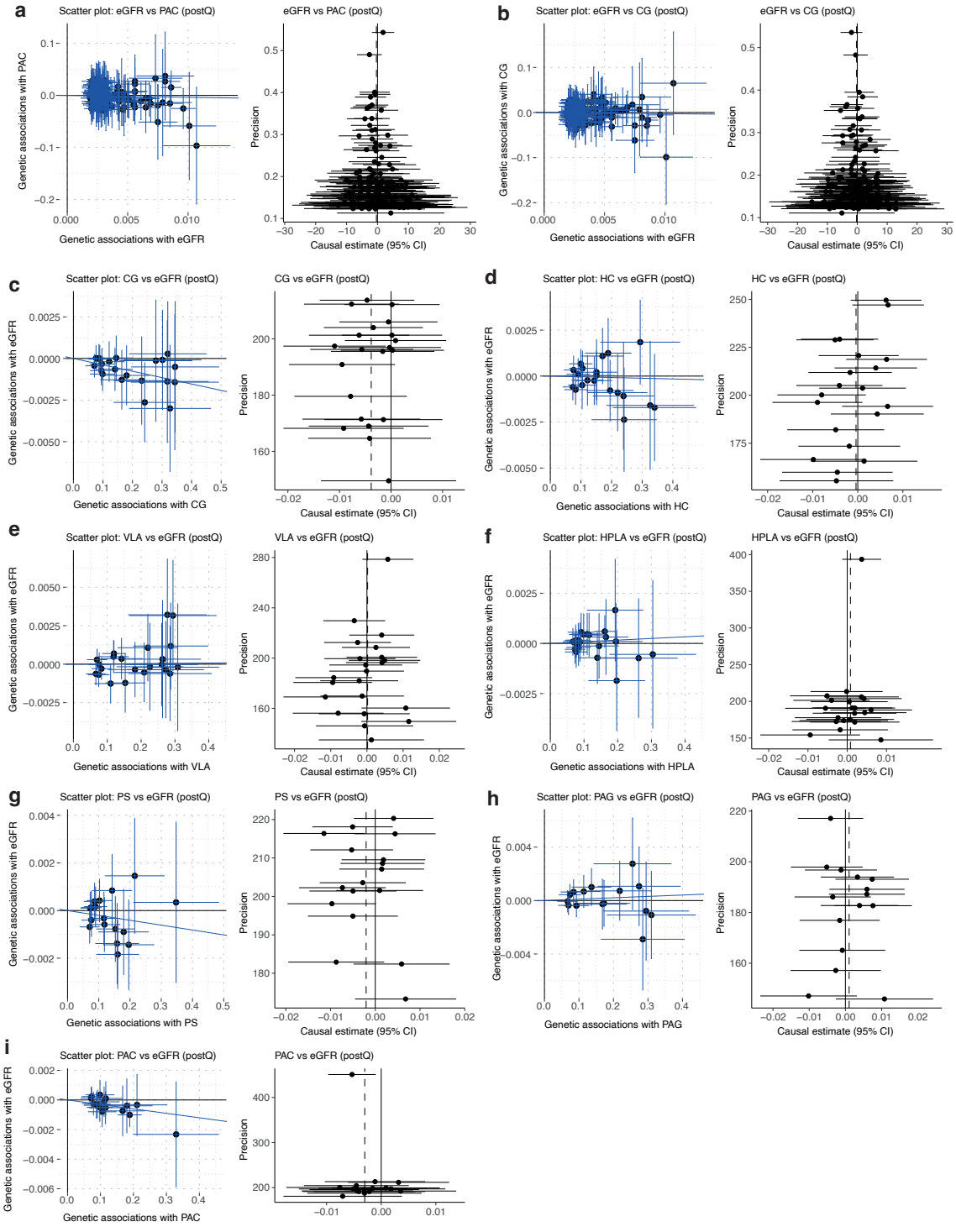
Supplementary Figure 5. Univariate Mendelian Randomization analyses of circulating metabolites on kidney function. Scatter and funnel plots revealing the associations among genetic variants associated with individual key metabolites of the phenylalanine-tyrosine pathway as exposures and eGFR as an outcome

(Supplementary Table 16). Genetic associations for eGFR were derived from CKDGen and for metabolites from CLSA and EGEA (please see methods). PCS, 4-cresyl sulfate; PCG, 4-cresyl glucuronide; PAA, phenylacetate; PAG, phenylacetylglutamine; PAC, phenylacetylcarnitine; HC, 3-phenylpropionate; CG, cinnamoylglycine; PS, phenol sulfate; HPLA, 3-(4-hydroxyphenyl)-lactate; VLA, vanillactate; proANP, pro-atrial natriuretic peptide; eGFR, estimated glomerular filtration ratio calculated using CKDepi-creatinine equation.



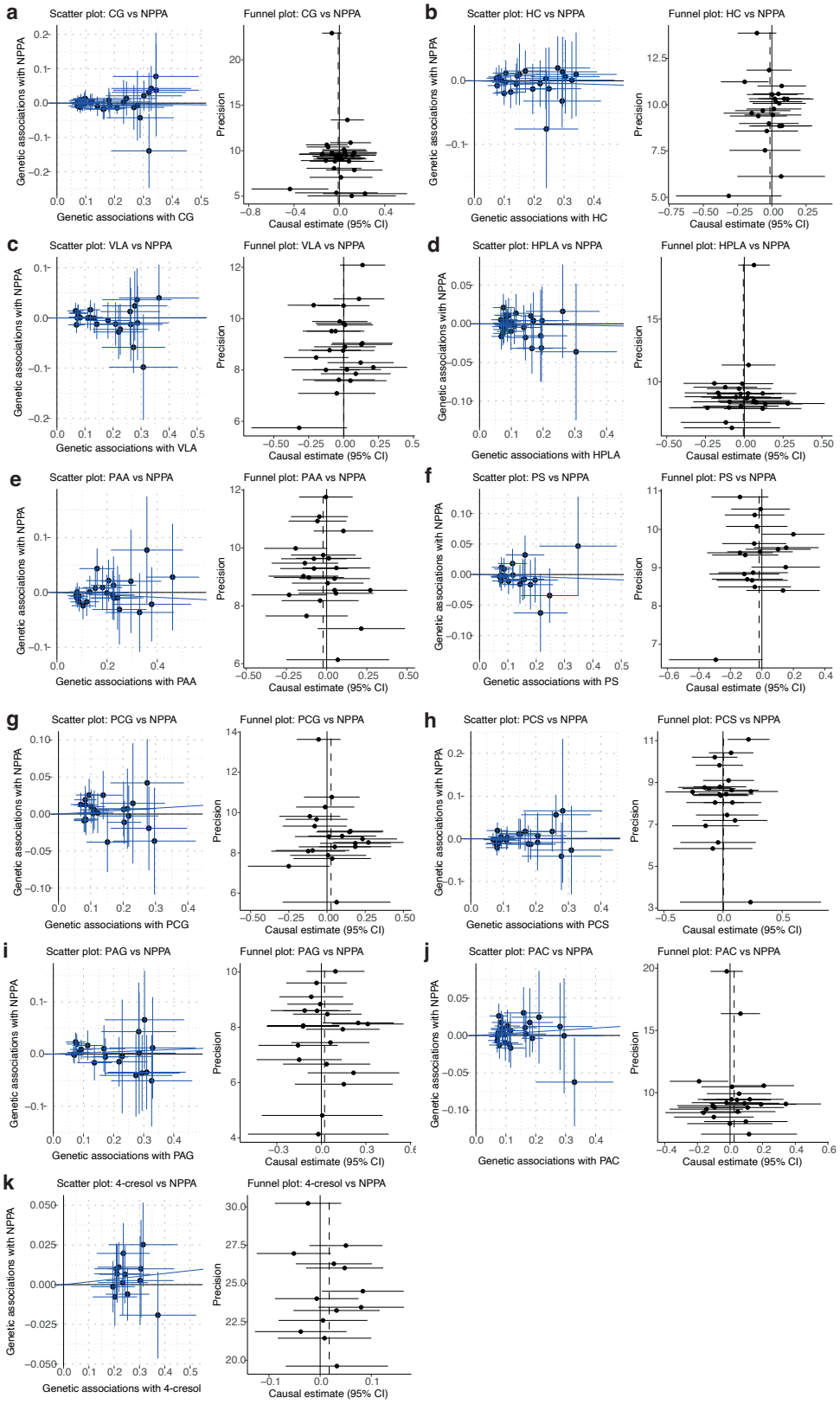
Supplementary Figure 6. Univariate Mendelian Randomization analyses of kidney function on circulating metabolites. Scatter and funnel plots revealing the

associations among genetic variants associated with eGFR as an exposure and key metabolites of the phenylalanine-tyrosine pathway as outcomes (Supplementary Table 17). Genetic associations for eGFR were derived from CKDGen and for metabolites from CLSA and EGEA (please see methods). PCS, 4-cresyl sulfate; PCG, 4-cresyl glucuronide; PAA, phenylacetate; PAG, phenylacetylglutamine; PAC, phenylacetylcarnitine; HC, 3-phenylpropionate; CG, cinnamoylglycine; PS, phenol sulfate; HPLA, 3-(4-hydroxyphenyl)-lactate; VLA, vanillactate; proANP, pro-atrial natriuretic peptide; eGFR, estimated glomerular filtration ratio calculated using CKDepi-creatinine equation.



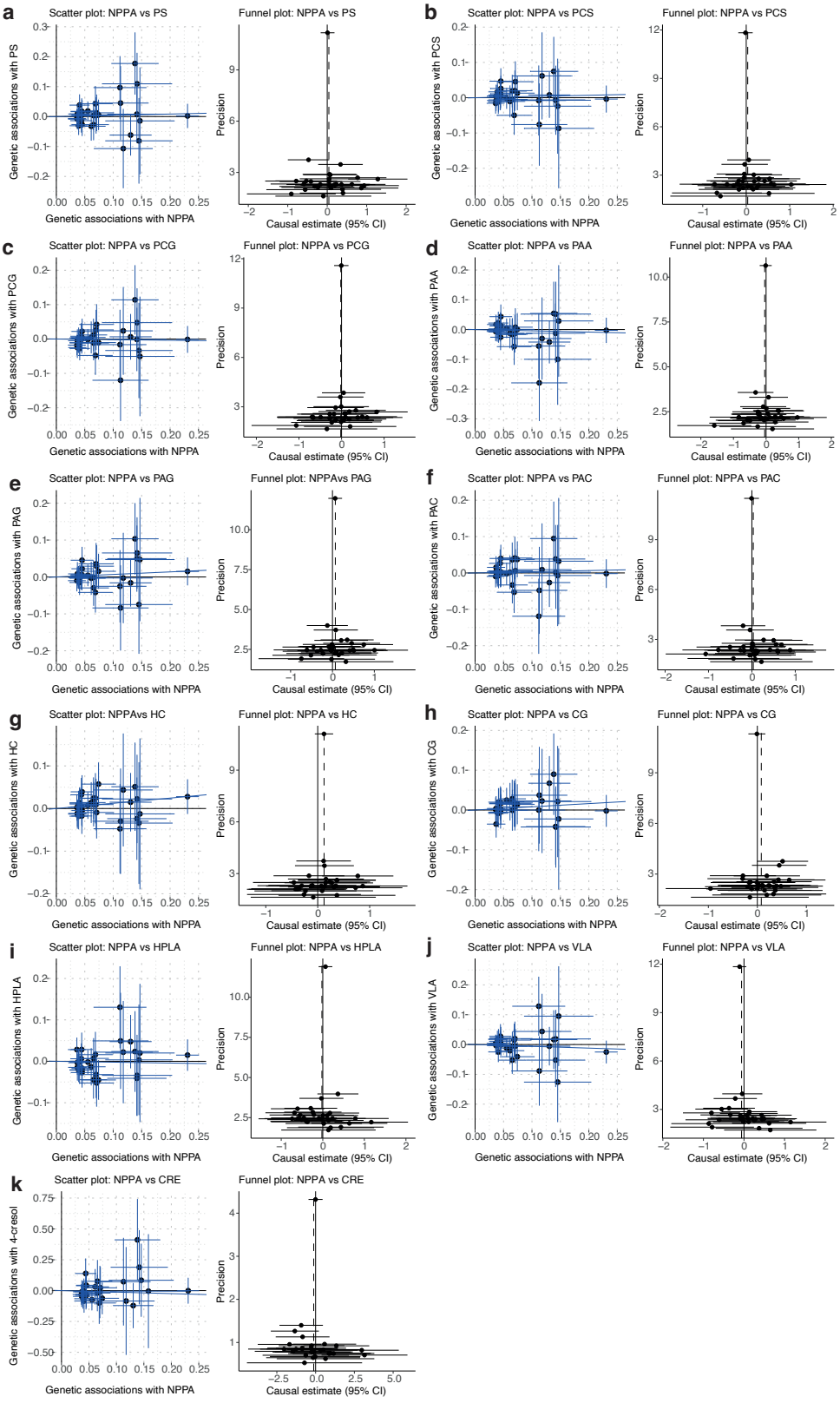
Supplementary Figure 7. Bi-directional univariate Mendelian Randomization analyses of circulating metabolites on kidney function post outlier correction. Scatter and funnel plots revealing the associations among genetic variants associated with individual key metabolites and eGFR post Cochrane's Q based outlier removal. Genetic associations for eGFR were derived from CKDGen and for metabolites from

CLSA and EGEA (please see methods). Only associations exhibiting significant heterogeneity (as measured by Cochrane's $Q \leq 0.05$ and I^2 statistics ≥ 25%) subjected to Q-based outlier corrections are shown here (Supplementary Tables 16,17). PCS, 4-cresyl sulfate; PCG, 4-cresyl glucuronide; PAA, phenylacetate; PAG, phenylacetylglutamine; PAC, phenylacetylcarnitine; HC, 3-phenylpropionate; CG, cinnamoylglycine; PS, phenol sulfate; HPLA, 3-(4-hydroxyphenyl)-lactate; VLA, vanillactate; proANP, pro-atrial natriuretic peptide; eGFR, estimated glomerular filtration ratio calculated using CKDepi-creatinine equation.



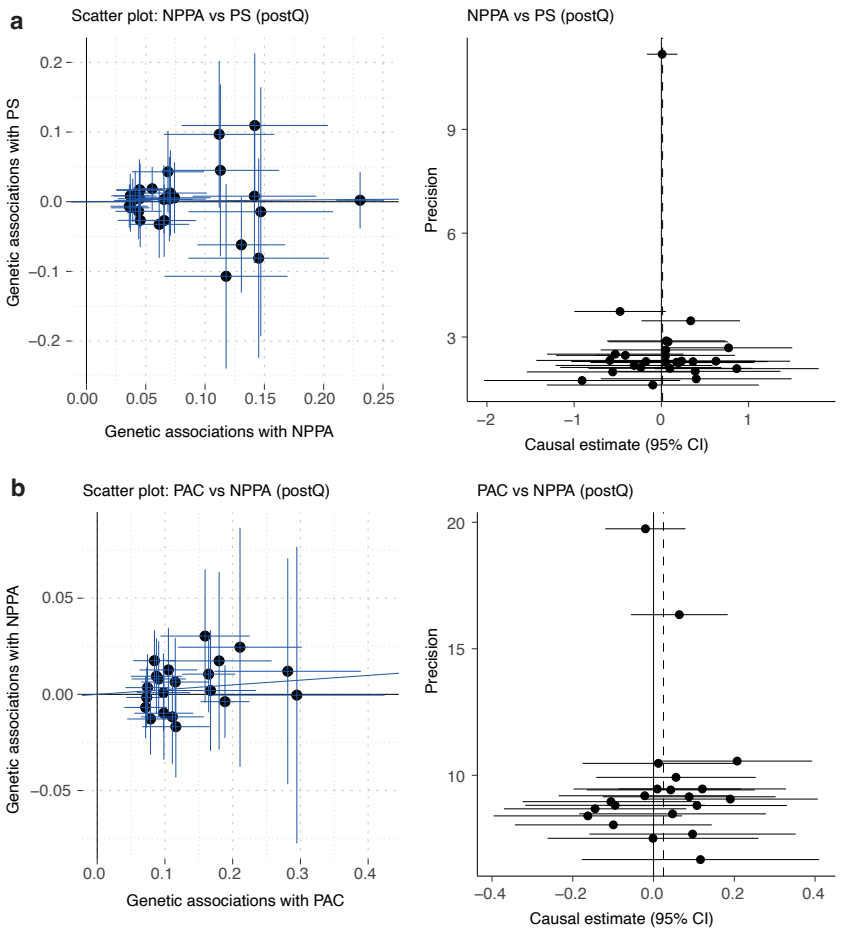
Supplementary Figure 8. Univariate Mendelian Randomization analyses of circulating metabolites on NPPA. Scatter and funnel plots revealing the associations

among genetic variants associated with individual key metabolites of the phenylalanine-tyrosine pathway as exposures and NPPA as an outcome (Supplementary Table 18). Genetic associations for NPPA were derived from deCODE and for metabolites from CLSA and EGEA (please see methods). PCS, 4-cresyl sulfate; PCG, 4-cresyl glucuronide; PAA, phenylacetate; PAG, phenylacetylglutamine; PAC, phenylacetylcarnitine; HC, 3-phenylpropionate; CG, cinnamoylglycine; PS, phenol sulfate; HPLA, 3-(4-hydroxyphenyl)-lactate; VLA, vanillactate; NPPA, Somascan based natriuretic peptide A relative quantification.

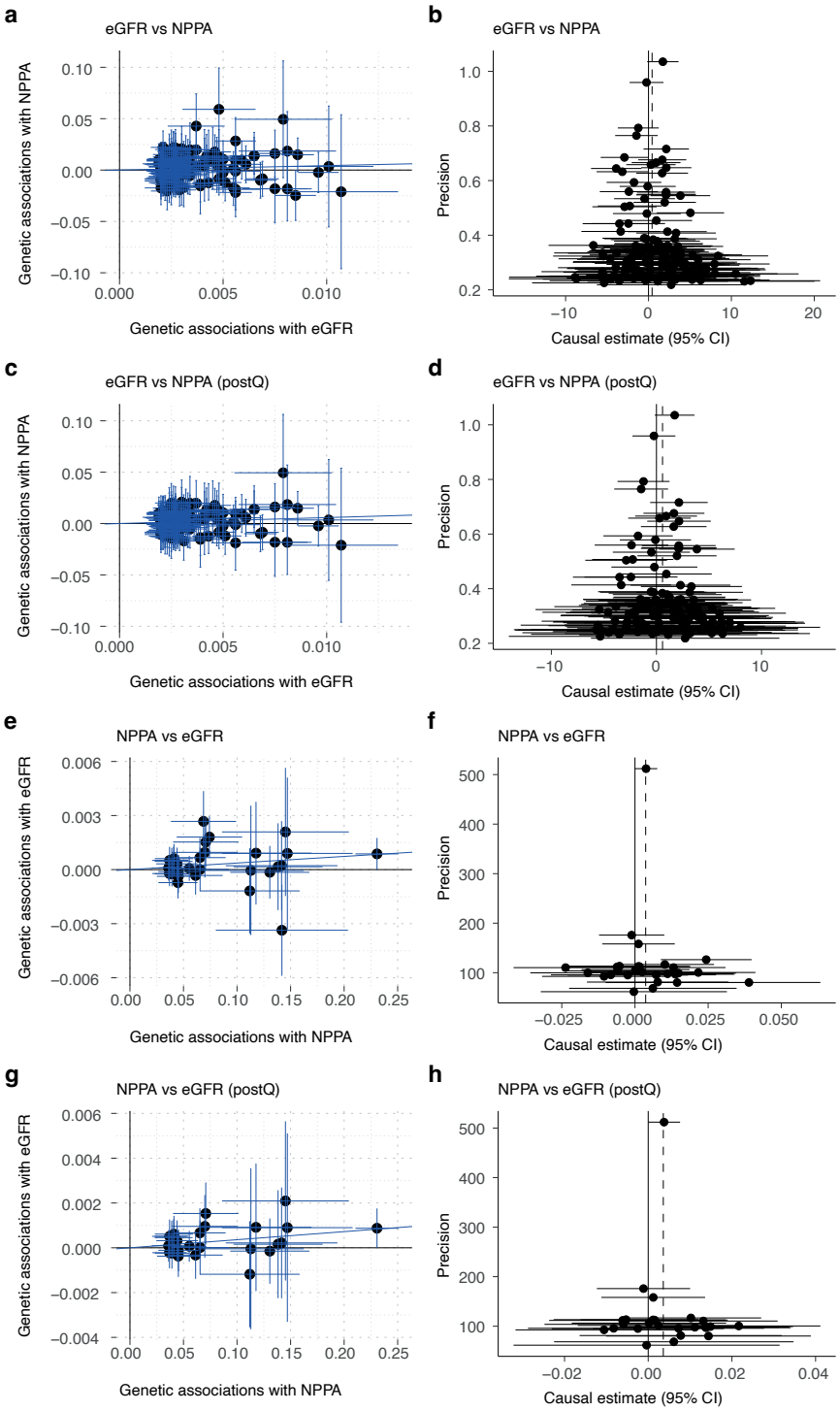


Supplementary Figure 9. Univariate Mendelian Randomization analyses of NPPA on circulating metabolites. Scatter and funnel plots revealing the associations

among genetic variants associated with NPPA as exposure and key metabolites of the phenylalanine-tyrosine pathway as outcomes (Supplementary Table 19). Genetic associations for NPPA were derived from deCODE and for metabolites from CLSA and EGEA (please see methods). PCS, 4-cresyl sulfate; PCG, 4-cresyl glucuronide; PAA, phenylacetate; PAG, phenylacetylglutamine; PAC, phenylacetylcarnitine; HC, 3-phenylpropionate; CG, cinnamoylglycine; PS, phenol sulfate; HPLA, 3-(4-hydroxyphenyl)-lactate; VLA, vanillactate; NPPA, Somascan based natriuretic peptide A relative quantification.



Supplementary Figure 10. Bi-directional univariate Mendelian Randomization analyses of circulating metabolites on NPPA post outlier correction. Scatter and funnel plots revealing the associations among genetic variants associated with individual key metabolites and NPPA post Cochrane's Q based outlier removal. Genetic associations for NPPA were derived from deCODE and for metabolites from CLSA and EGEA (please see methods). Only associations exhibiting significant heterogeneity (as measured by Cochrane's $Q \leq 0.05$ and I^2 statistics ≥ 25%) subjected to Q-based outlier corrections are shown here (Supplementary Tables 18, 19). PCS, 4-cresyl sulfate; PCG, 4-cresyl glucuronide; PAA, phenylacetate; PAG, phenylacetylglutamine; PAC, phenylacetylcarnitine; HC, 3-phenylpropionate; CG, cinnamoylglycine; PS, phenol sulfate; HPLA, 3-(4-hydroxyphenyl)-lactate; VLA, vanillactate; NPPA, Somascan based natriuretic peptide A relative quantification.



Supplementary Figure 11. Bi-directional univariate Mendelian Randomization analyses of eGFR-NPPA axis. Scatter and funnel plots revealing the associations among genetic variants associated with eGFR and NPPA pre- and post- Cochrane's Q based outlier removal (Supplementary Table 20). Genetic associations for eGFR and NPPA were derived from CKDGen and deCODE, respectively (please see

methods). eGFR, estimated glomerular filtration ratio calculated using CKDepi-creatinine equation; NPPA, Somascan based natriuretic peptide A relative quantification.