

Lipidomic analyses of large cohort studies define the role of lipid metabolism in bridging diet and cardio-metabolic health

Habtamu B Beyene^{1,2,3}, Tingting Wang^{1,2,3}, Michelle Cinel¹, Natalie A Mellett¹, Thy Duong¹, Matilda van Buuren-Milne¹, Alexandra N. Faulkner¹, Jingqin Wu^{1,3}, Gavriel Olshansky^{1,3}, Jonathan E Shaw^{1,4}, Dianna J Magliano^{1,4}, Melissa C. Southy^{4,5,6}, Roger L. Milne^{4,5,6}, Allison M. Hodge^{5,6}, Corey Giles^{1,2,3}, Kevin Huynh^{1,2,3,^,*}, Peter J Meikle^{1,2,3,4,^,*}

¹Baker Heart and Diabetes Institute, Melbourne, VIC, Australia.

²Baker Department of Cardiovascular Research, Translation and Implementation, La Trobe University, Melbourne, VIC, Australia.

³Baker Department of Cardiometabolic Health, Melbourne University, Melbourne, VIC, Australia.

⁴Precision Medicine, School of Clinical Sciences at Monash Health, Monash University, Clayton, VIC, Australia.

⁵Cancer Epidemiology Division, Cancer Council Victoria, Melbourne, VIC, Australia.

⁶Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global Health, The University of Melbourne, VIC, Australia

*Co-corresponding authors

^These authors jointly supervised this work

Peter J Meikle PhD

Head, Metabolomics Laboratory

Baker Heart and Diabetes Institute, 75 Commercial Road, Melbourne VIC 3004

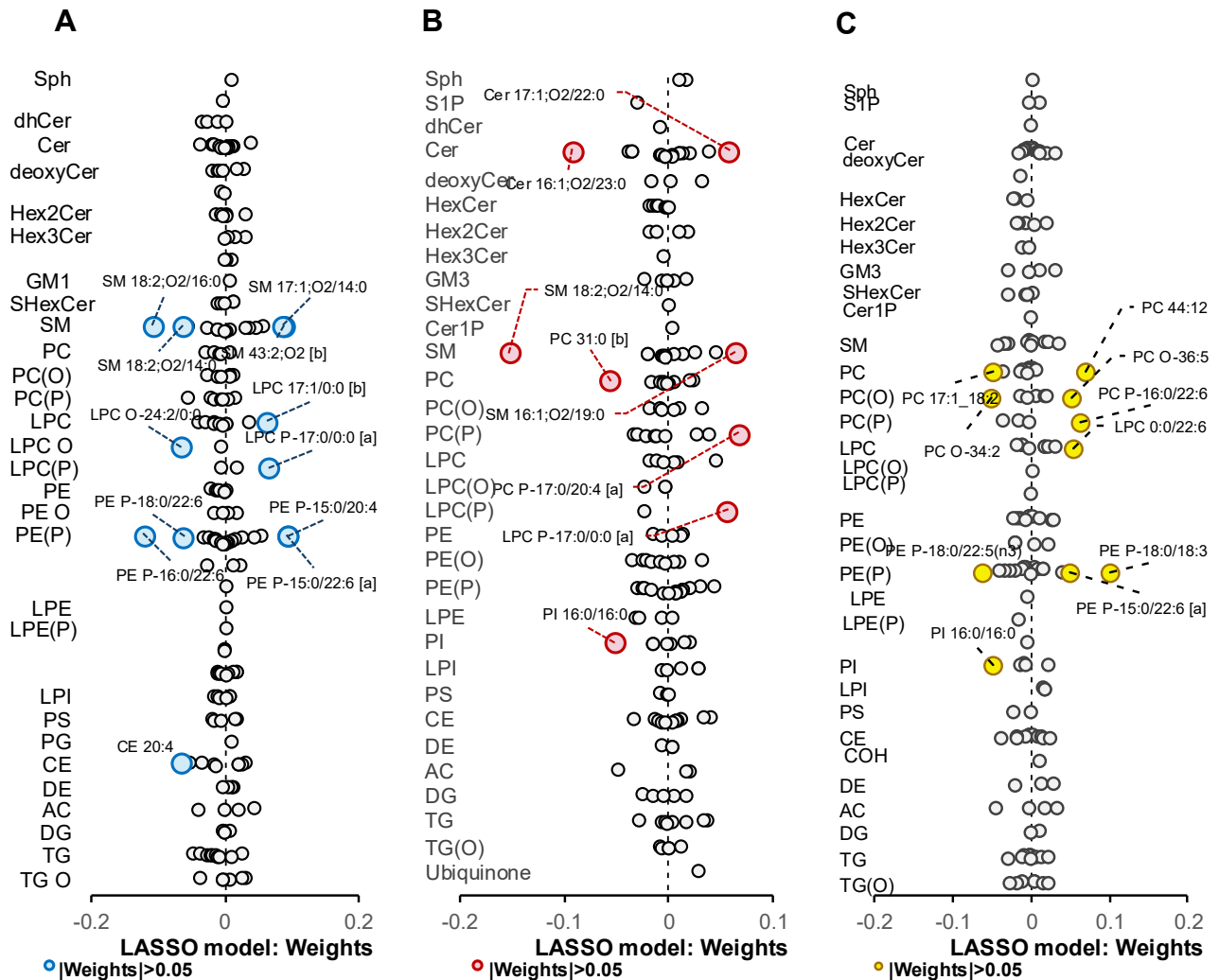
T +61 (0)3 8532 1770; E: peter.meikle@baker.edu.au

Kevin Huynh PhD

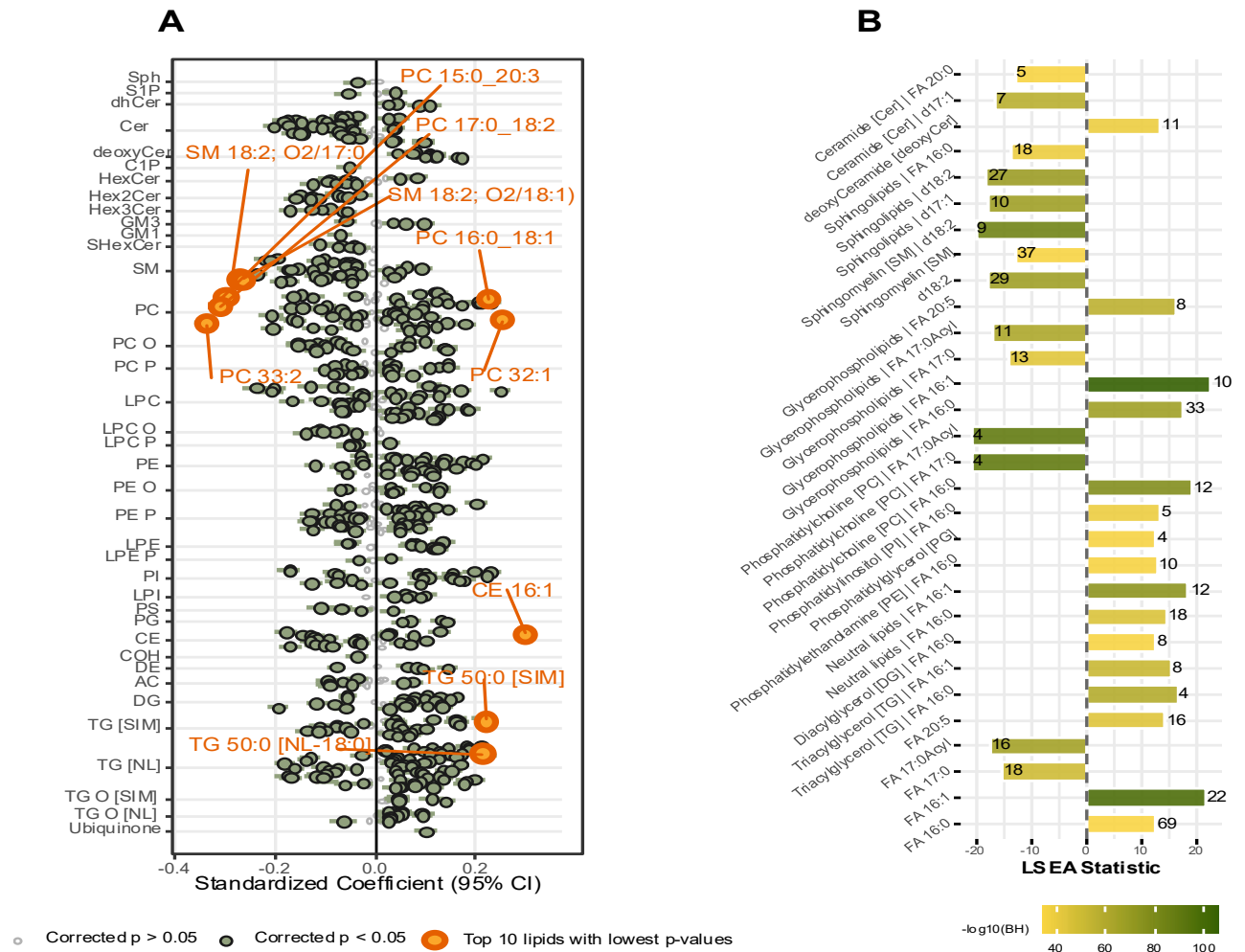
Group leader, Metabolomics Laboratory

Baker Heart and Diabetes Institute, 75 Commercial Road, Melbourne VIC 3004

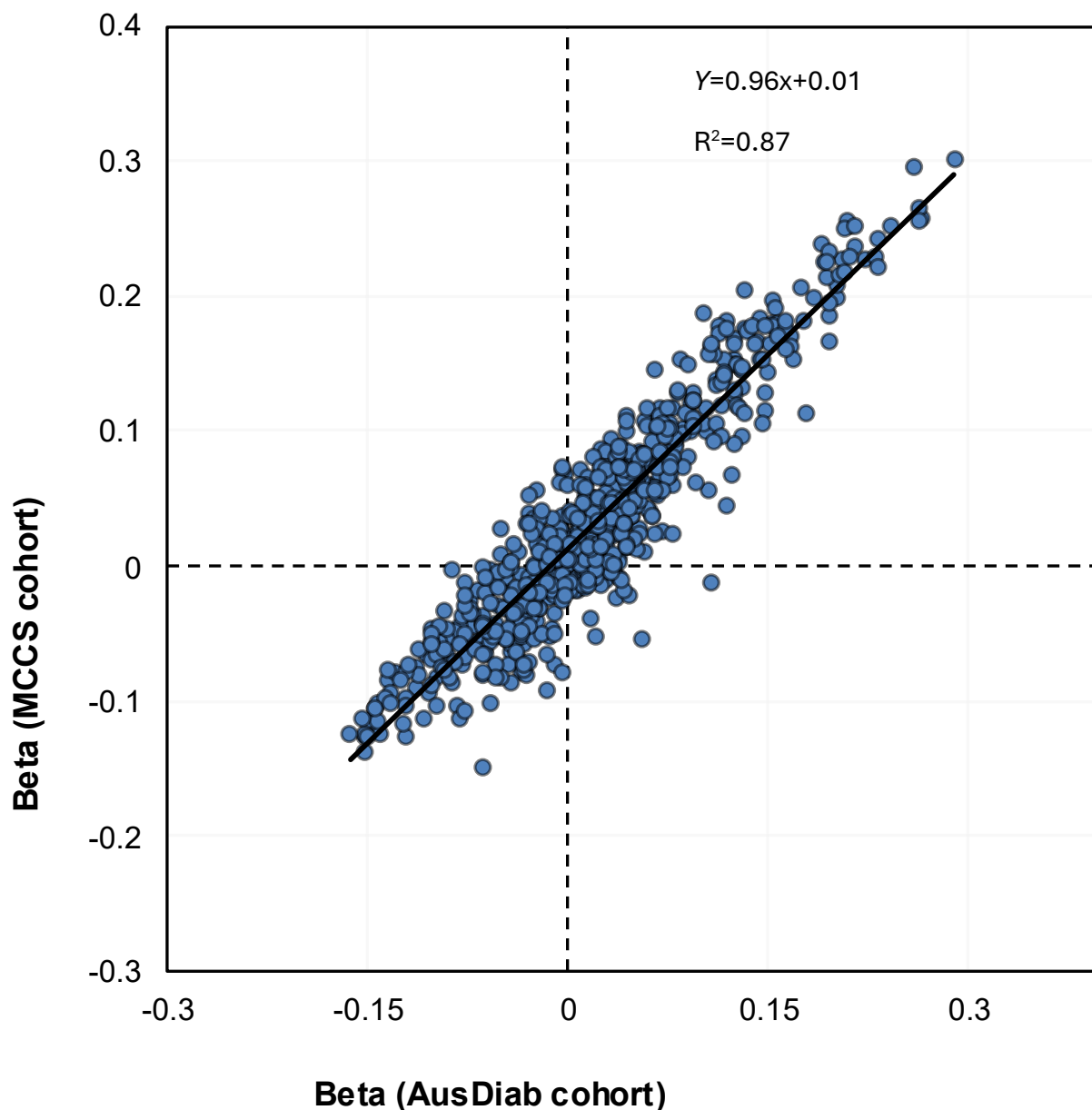
T +61 (0)3 8532 1770; E: kevin.huynh@baker.edu.au



Supplementary Fig. 1 | LASSO-selected lipid species and their coefficient weights for dietary intake models. LASSO regression models were fitted in the AusDiab cohort (n = 10,339 independent samples) to identify lipid species associated with dietary intake of (A) dairy, (B) red meat, and (C) fish. Each dot represents a lipid coefficient weight from the fitted model, with coloured circles highlighting the top features with an absolute coefficient magnitude > 0.05. Source data are provided in the Source Data file.

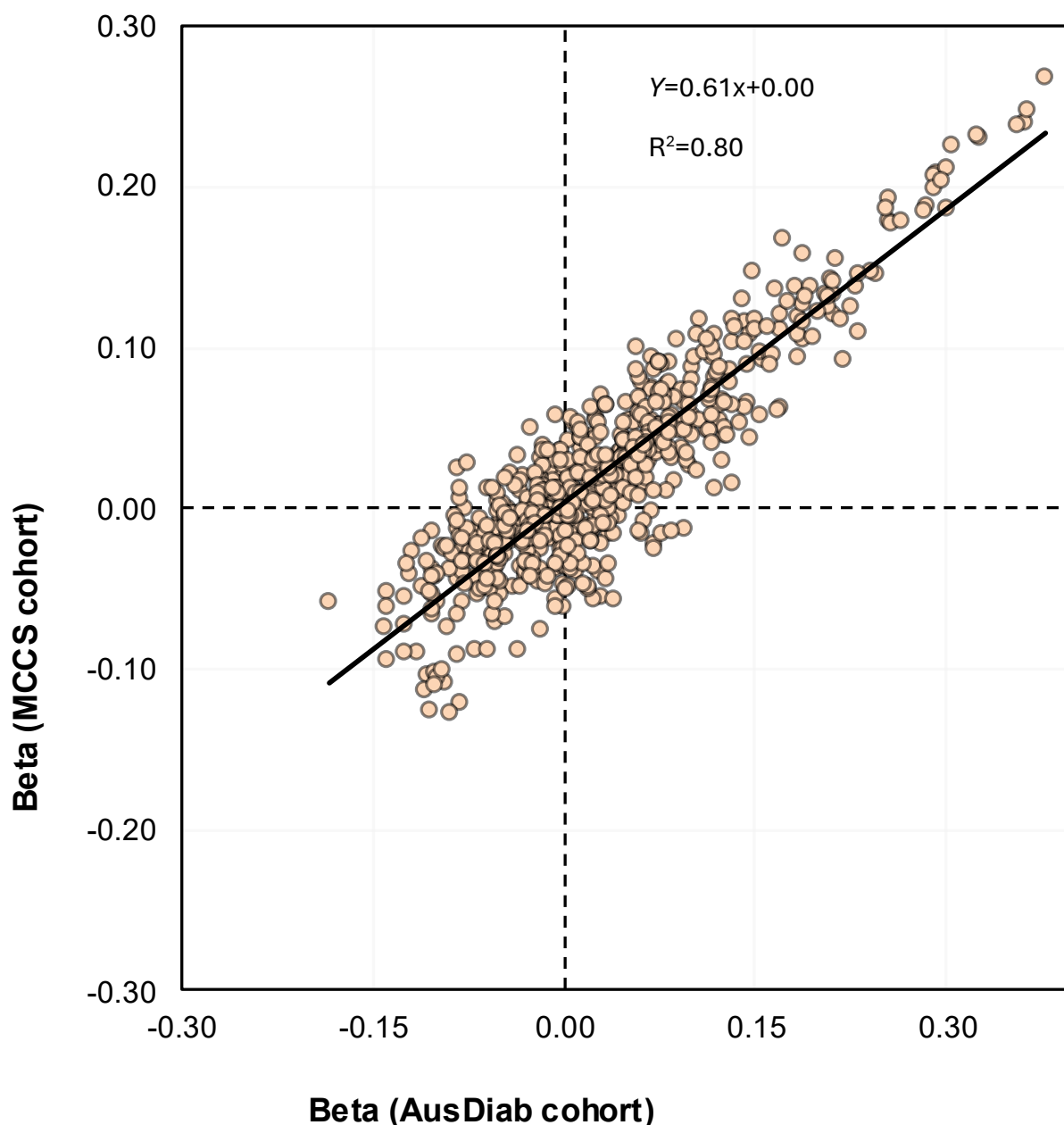


Supplementary Fig. 2 | Associations between alcohol intake and the plasma lipidome in the AusDiab cohort. Panel A shows associations with individual lipid species, and Panel B shows lipid set-level associations. Linear regression and lipid set enrichment analysis (LSEA) were performed with inverse-normal-transformed alcohol intake as the predictor and scaled lipid species concentrations or lipid sets (unit SD) as outcomes. Models were adjusted for age, sex, BMI, energy intake, clinical lipids (total cholesterol, HDL-C, and triglycerides), exercise, diabetes status, and smoking. Each data point represents a beta coefficient per 1-SD higher alcohol intake, with error bars indicating 95% confidence intervals. Two-sided p values are shown for each lipid species: open circles indicate $p > 0.05$, and dark green circles indicate $p < 0.05$ after Benjamini-Hochberg correction for multiple testing. Brown circles highlight the top 10 lipid species ranked by p value. Lipid class abbreviations are as follows: AC, acylcarnitine; CE, cholesteryl ester; Cer, ceramide; COH, cholesterol; DE, dehydrocholesterol; dhCer, dihydroceramide; DG, diacylglycerol; GM1, GM1 ganglioside; GM3, GM3 ganglioside; HexCer, monohexosylceramide; Hex2Cer, dihexosylceramide; Hex3Cer, trihexosylceramide; LPC, lysophosphatidylcholine; LPC O, lysoalkylphosphatidylcholine; LPC P, lysoalkenylphosphatidylcholine; LPE, lysophosphatidylethanolamine; LPE P, lysoalkenylphosphatidylethanolamine; LPI, lysophosphatidylinositol; PC, phosphatidylcholine; PC O, alkylphosphatidylcholine; PC P, alkenylphosphatidylcholine; PE, phosphatidylethanolamine; PE O, alkylphosphatidylethanolamine; PE P, alkenylphosphatidylethanolamine; PG, phosphatidylglycerol; PI, phosphatidylinositol; PS, phosphatidylserine; SHexCer, sulfatide; SM, sphingomyelin; TG, triacylglycerol; TG O, alkyl-diacylglycerol. NL, neutral loss; SIM, single ion monitoring. Letters (a), (b), and (c) denote chromatographic elution order; for example, PE P-17:0/22:6 (a) elutes prior to PE P-17:0/22:6 (b). Source data are provided in the Source Data file.



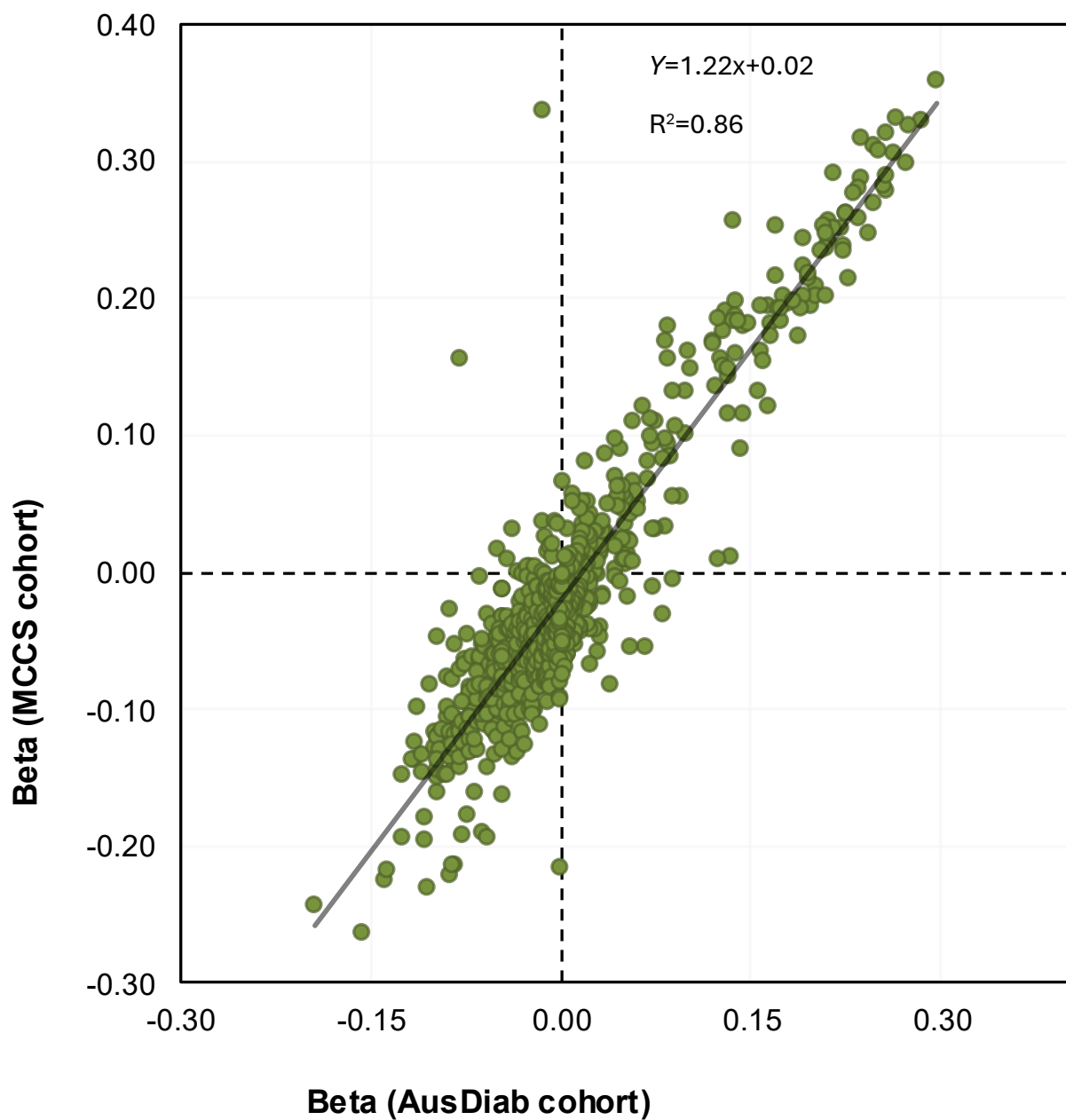
Supplementary Fig. 3 | Replication of dairy–lipidome associations across cohorts.

Scatter plots show the correlation of regression coefficients for dairy intake–associated lipid species between cohorts, with AusDiab estimates on the x-axis and Melbourne Collaborative Cohort Study (MCCS) estimates on the y-axis. Coefficients were obtained from linear regression models fitted independently in the AusDiab cohort (n = 10,339 samples) and the MCCS (n = 2,996 samples). AusDiab denotes the Australian Diabetes, Obesity and Lifestyle Study. Source data are provided in the Source Data file.

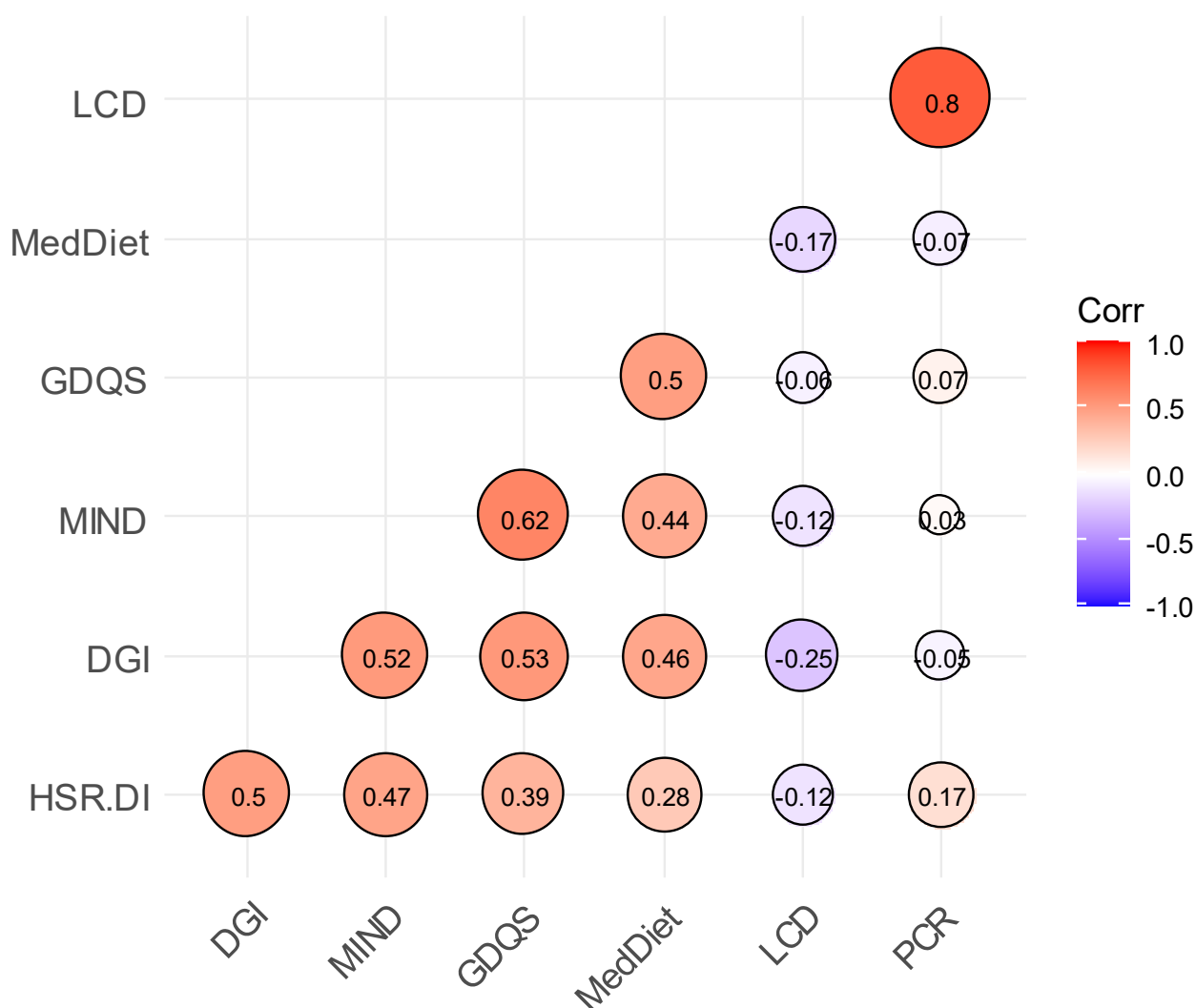


Supplementary Fig. 4 | Replication of red meat intake–lipidome associations.

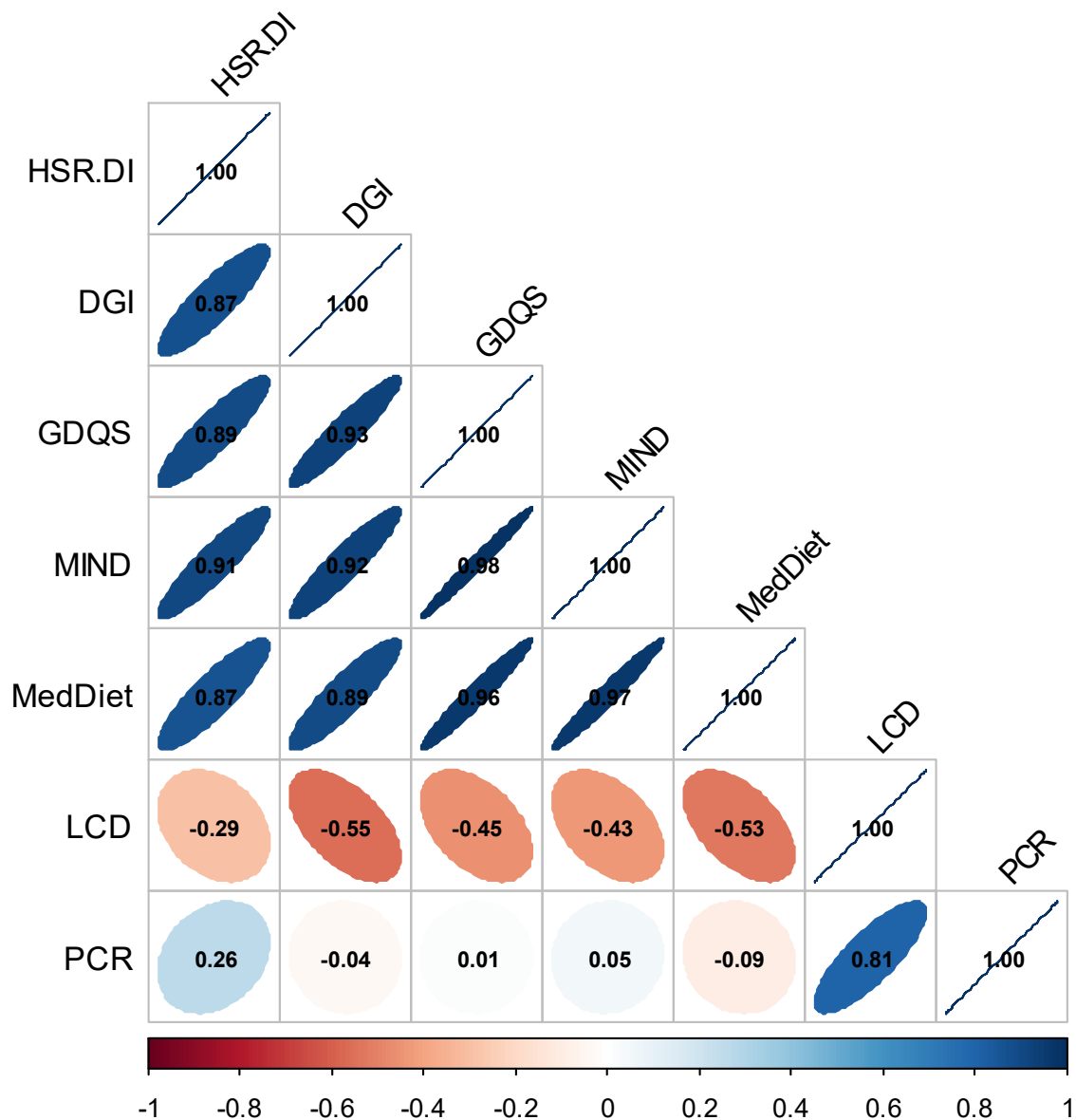
Correlation of beta coefficients between cohorts, with each data point representing estimates from AusDiab (x-axis) and MCCS (y-axis). Coefficients were derived from linear regression models in the AusDiab cohort (n = 10,339 samples) and the Melbourne Collaborative Cohort Study (MCCS; n = 2,996 samples). AusDiab, Australian Diabetes, Obesity and Lifestyle Study. Source data are provided in the Source Data file.



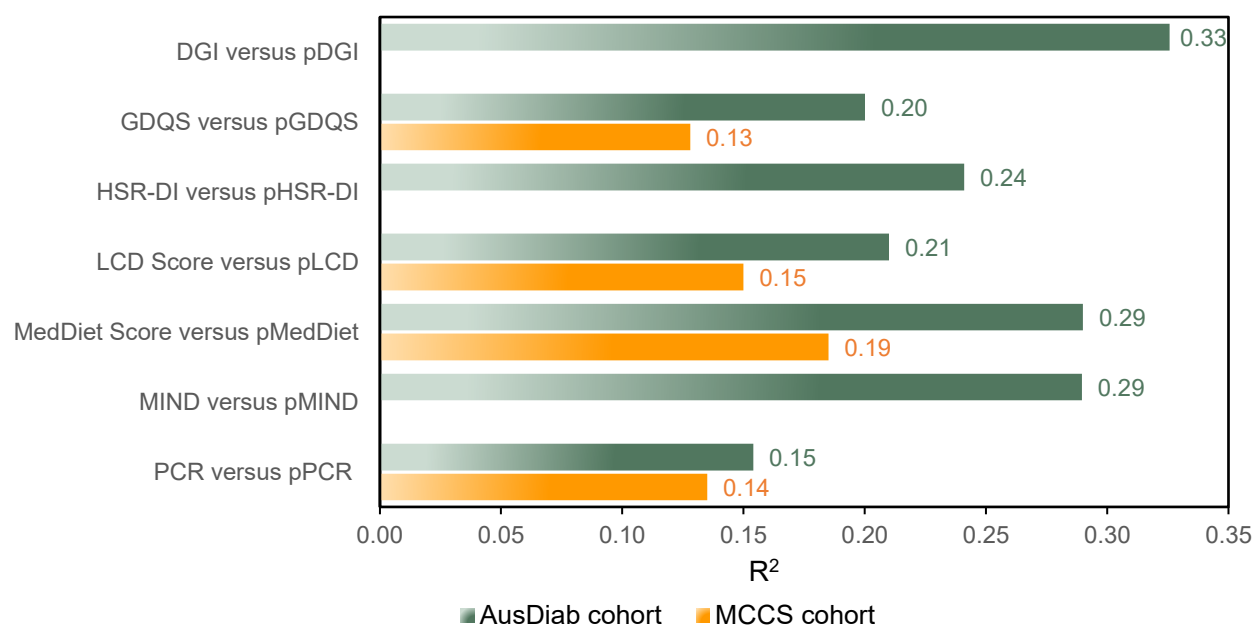
Supplementary Fig. 5 | Replication of fish intake–lipidome associations. Correlation of beta coefficients between cohorts, with each data point representing estimates from AusDiab (x-axis) and MCCS (y-axis). Coefficients were derived from linear regression models in the AusDiab cohort (n = 10,339 samples) and the Melbourne Collaborative Cohort Study (MCCS; n = 2,996 samples). AusDiab, Australian Diabetes, Obesity and Lifestyle Study. Source data are provided in the Source Data file.



Supplementary Fig. 6 | Correlations between FFQ-derived diet quality scores. Pairwise Pearson correlation coefficients between diet quality scores derived from food frequency questionnaire (FFQ) data in the AusDiab cohort (n = 10,339 participants). Values shown within each circle indicate correlation coefficients, with circle size and colour intensity reflecting correlation strength. Colours range from blue (negative correlations) to red (positive correlations), and circle size increases with increasing correlation magnitude. DGI, Australian Dietary Guideline Index; GDQS, Global Diet Quality Score; HSR, Health Star Rating–Dietary Index; LCD, Low-Carbohydrate Diet; MedDiet, Mediterranean Diet Score; MIND, Mediterranean–DASH (Dietary Approaches to Stop Hypertension) Intervention for Neurodegenerative Delay; PCR, Protein-to-Carbohydrate Ratio.



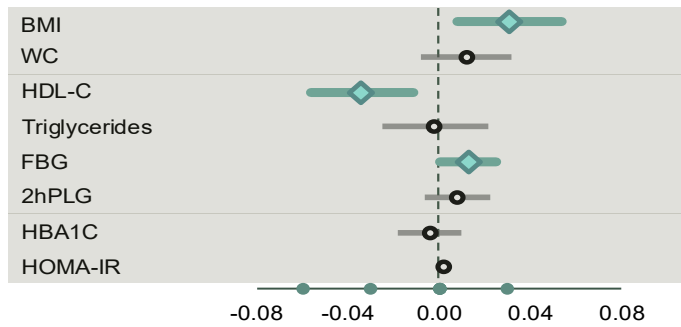
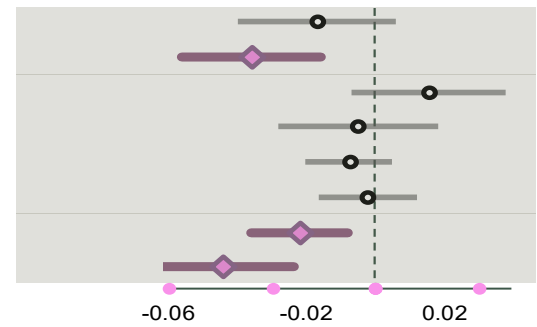
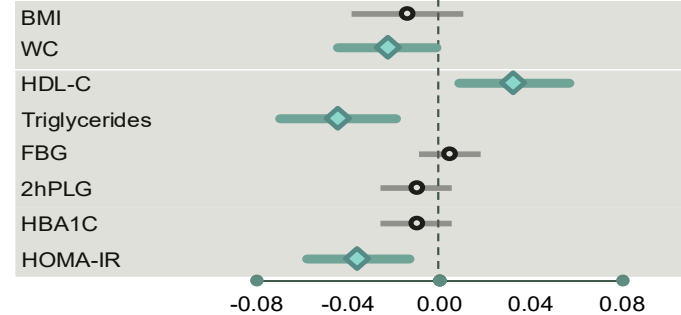
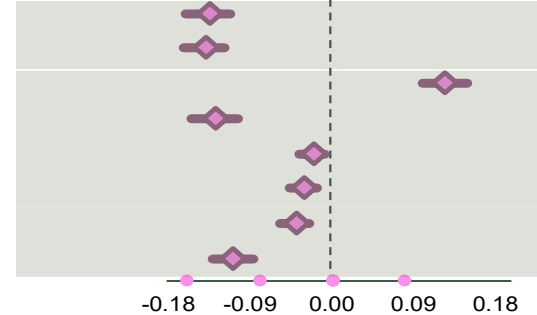
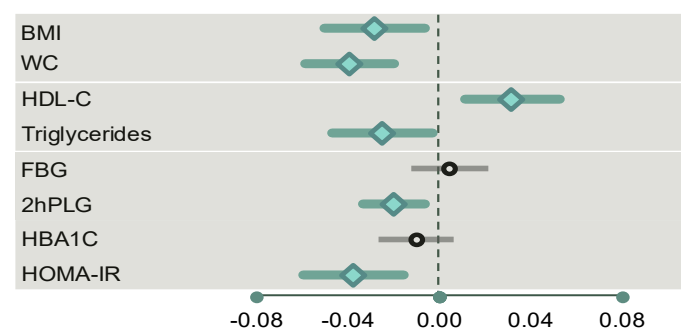
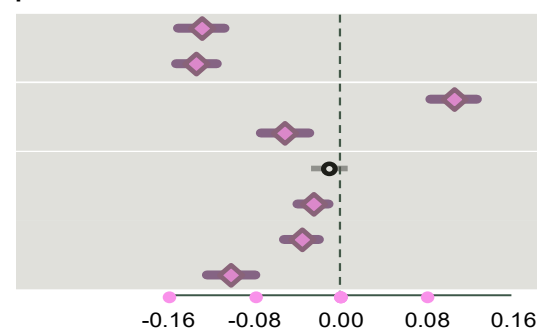
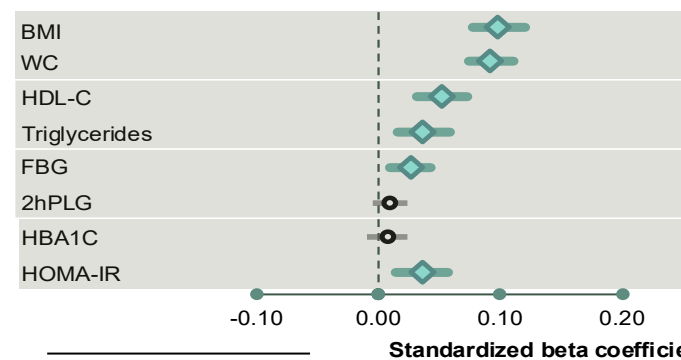
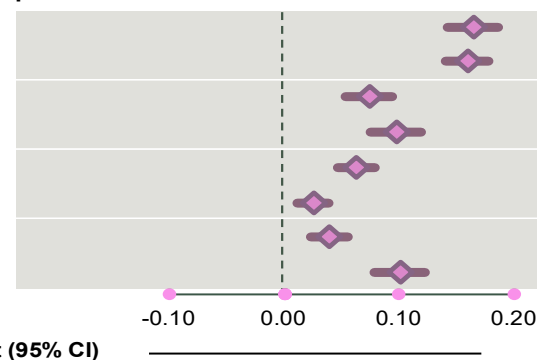
Supplementary Fig. 7 | Correlation of lipidomic association patterns across diet quality scores. Pairwise correlations of lipid species-specific beta coefficients derived from regression models of different diet quality scores in the AusDiab cohort (n = 10,339 participants). For each lipid species, beta coefficients were correlated across diet quality scores. Values within each ellipse indicate Pearson's correlation coefficients. Ellipse colour and shape reflect correlation direction and strength, with colours ranging from blue (positive correlations) to dark red (negative correlations), and ellipses transitioning from wider to narrower shapes to indicate weaker to stronger correlations, respectively. DGI, Australian Dietary Guideline Index; GDQS, Global Diet Quality Score; HSR, Health Star Rating–Dietary Index; LCD, Low-Carbohydrate Diet; MedDiet, Mediterranean Diet Score; MIND, Mediterranean–DASH (Dietary Approaches to Stop Hypertension) Intervention for Neurodegenerative Delay; PCR, Protein-to-Carbohydrate Ratio.



Supplementary Fig. 8 | Prediction of diet quality scores using the plasma lipidome.

Correlation (R^2) between lipidome-predicted diet quality scores and FFQ-derived diet quality scores (x-axis) in the AusDiab (yellow bars) and MCCS (light green bars) cohorts. Data are based on $n = 9,430$ biologically independent samples in AusDiab and $n = 2,700$ in MCCS.

DGI, Australian Dietary Guideline Index; GDQS, Global Diet Quality Score; HSR, Health Star Rating–Dietary Index; LCD, Low-Carbohydrate Diet; MedDiet, Mediterranean Diet Score; MIND, Mediterranean–DASH (Dietary Approaches to Stop Hypertension) Intervention for Neurodegenerative Delay; PCR, Protein-to-Carbohydrate Ratio. Diet scores with the prefix “p” denote lipidome-predicted diet quality scores.

HSR-DI**pHSR-DI****GDQS****pGDQS****MIND****pMIND****LCD score****pLCD score**

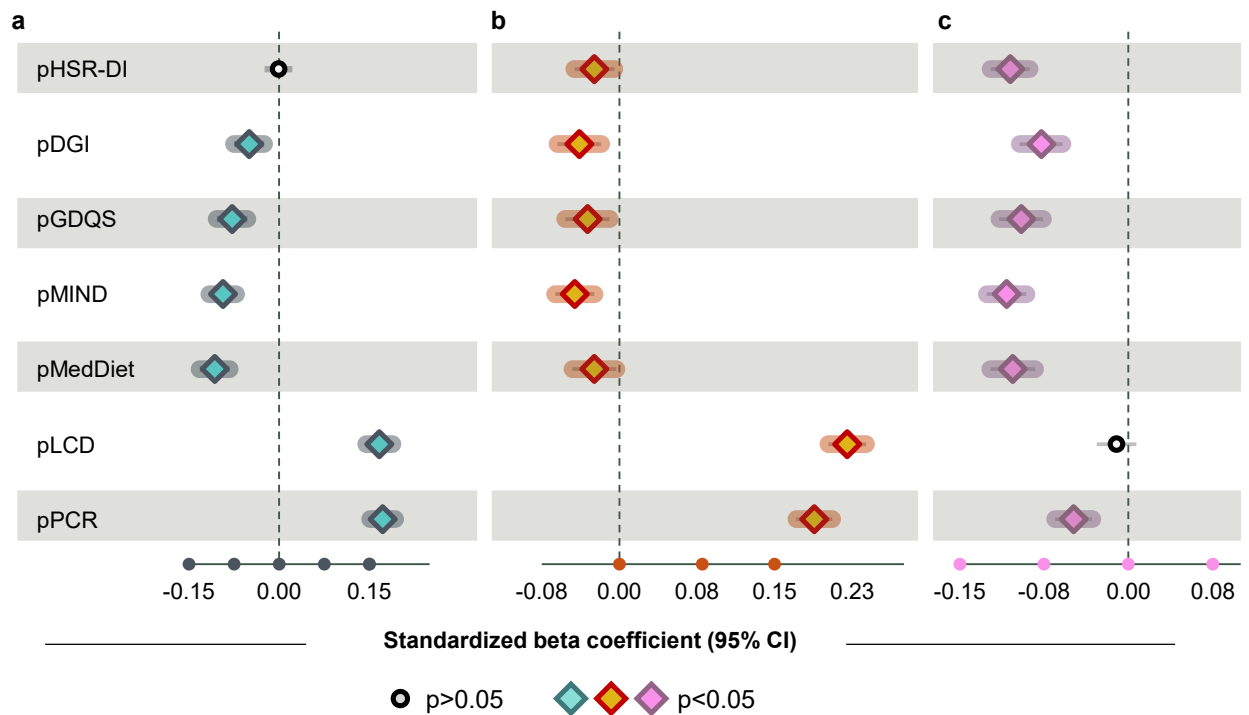
Standardized beta coefficient (95% CI)

○ p > 0.05

◆

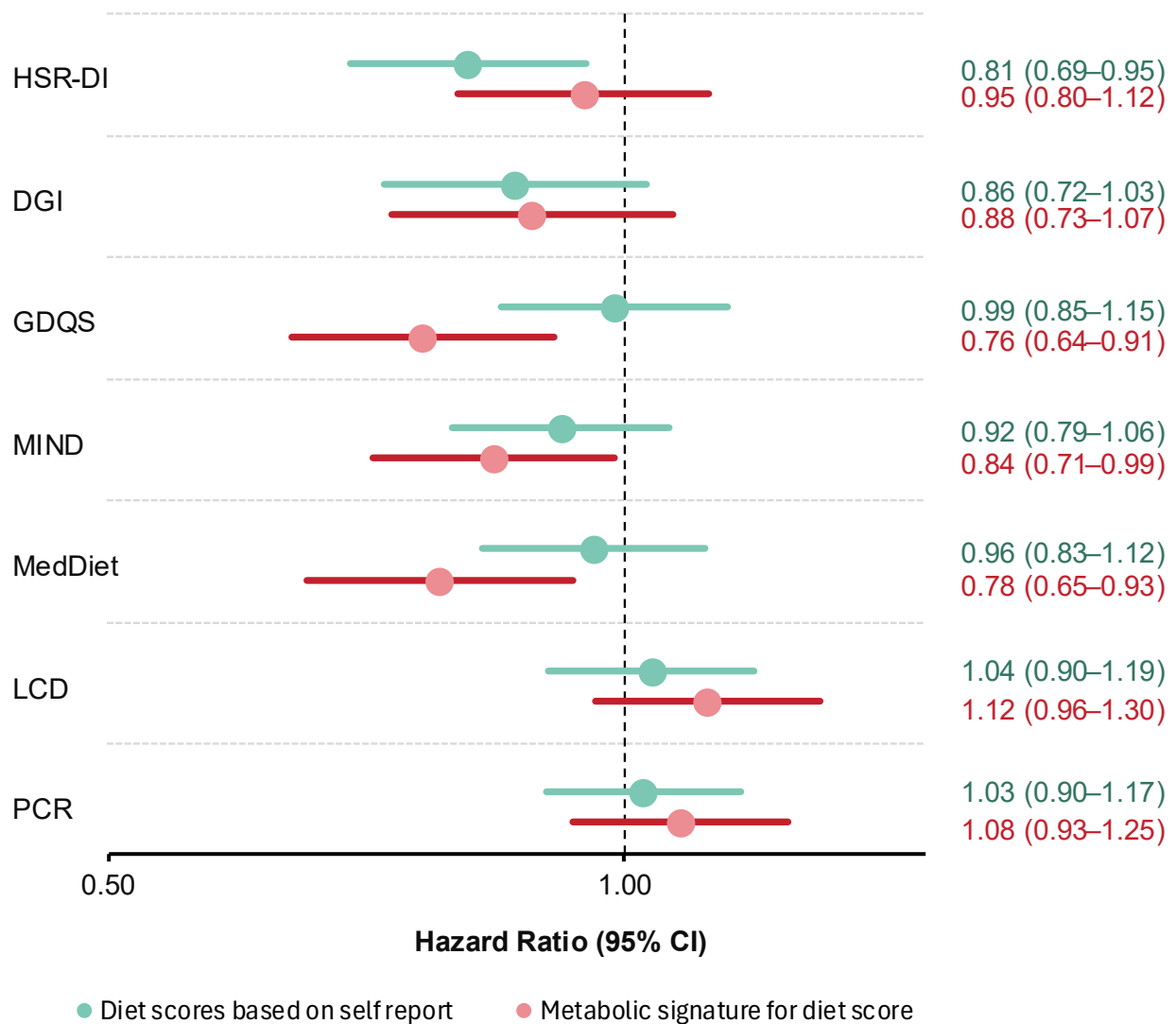
◆ p < 0.05

Supplementary Fig. 9 | Associations of FFQ-derived and lipidome-predicted diet quality scores with conventional cardiometabolic health biomarkers. Each circle or diamond in the forest plots represents standardized beta coefficients. Error bars denote 95% confidence intervals (CIs) derived from multiple linear regression models assessing associations between metabolic risk scores and diet quality scores, adjusted for age, sex, energy intake, smoking, physical activity, and diabetes status. Data are based on the AusDiab cohort (n = 10,339 participants). Open circles indicate p > 0.05, whereas coloured diamonds indicate two-sided p values < 0.05.



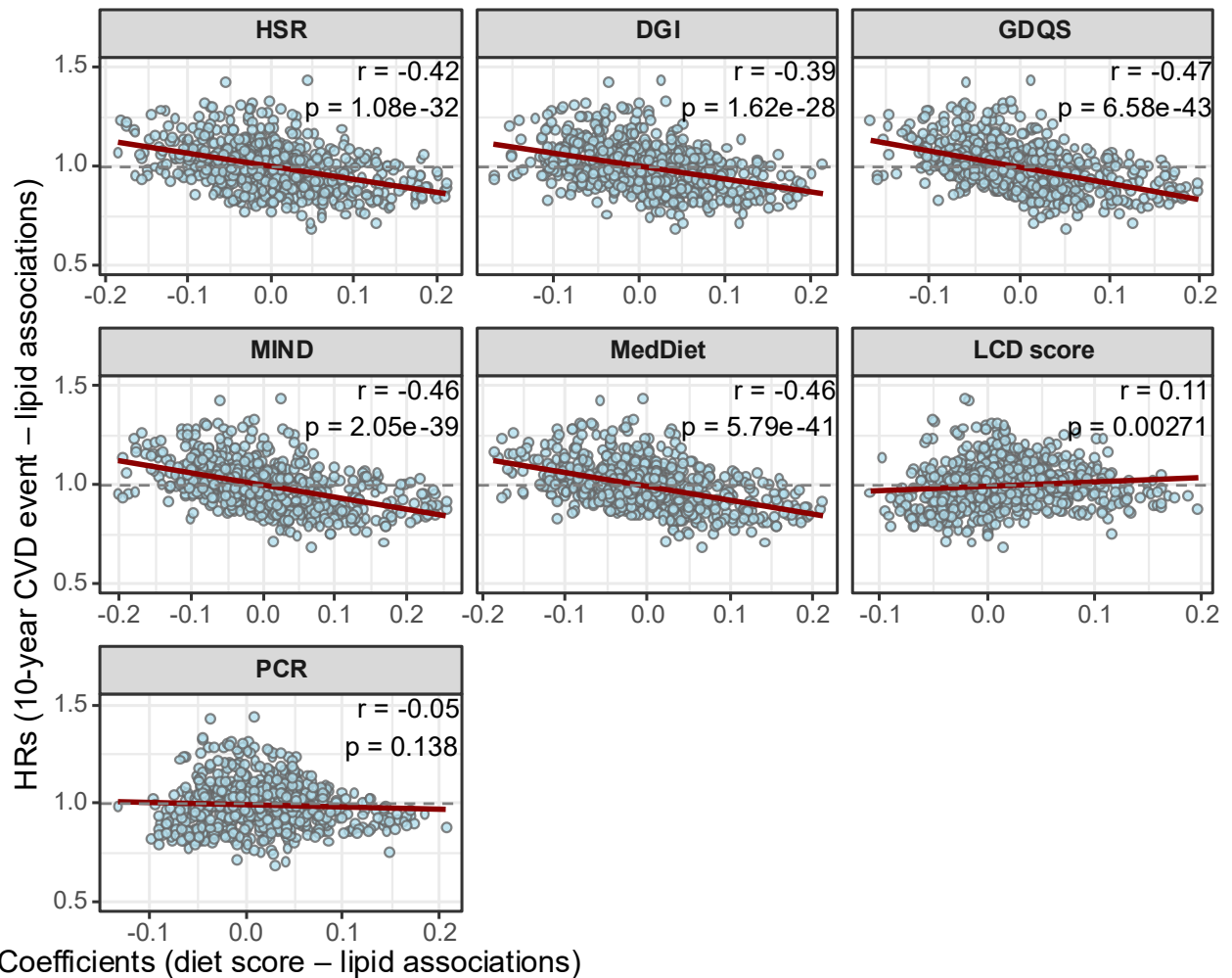
Supplementary Fig. 10 | Associations of lipidome-predicted metabolic signatures of diet quality scores with metabolic health: (a) metabolic BMI, (b) plasmalogen score, and (c) lipidomic risk score (LRS). Each circle or diamond in the forest plots represents standardized beta coefficients, with error bars indicating 95% confidence intervals (CIs). Estimates were derived from multiple linear regression models assessing associations between metabolic health measures and diet quality scores, adjusted for age, sex, BMI, clinical lipids (total cholesterol, HDL-C, triglycerides), smoking, physical activity, and diabetes status in the AusDiab cohort ($n = 10,339$ participants). Beta coefficients are expressed per 1-SD increment.

LRS, lipidomic risk score for cardiovascular disease risk. Open circles indicate $p > 0.05$, whereas blue, brown, and pink diamonds indicate two-sided $p < 0.05$. The prefix “p” denotes lipidome-predicted scores.



Supplementary Fig. 11 | Associations of FFQ-derived diet quality scores and their corresponding lipidome predicted metabolic signatures with CVD mortality. Each data point in the forest plots represents hazard ratios (HRs) per 1-SD increment, with error bars showing 95% confidence intervals (CIs) estimated using Cox proportional hazards regression. Models were adjusted for age, sex, BMI, clinical lipids (total cholesterol, HDL-C, triglycerides), smoking, physical activity, hypertension, lipid-lowering medication, energy intake, diabetes status, and diet scores. Data are based on the AusDiab cohort, with 264 CVD events and 8,000 controls over a 10-year follow-up. Green circles indicate estimates for FFQ/self-reported diet scores, and red circles indicate estimates for the corresponding metabolic signatures.

DGI, Australian Dietary Guideline Index; GDQS, Global Diet Quality Score; HSR, Health Star Rating–Dietary Index; LCD, Low-Carbohydrate Diet; MedDiet, Mediterranean Diet Score; MIND, Mediterranean–DASH (Dietary Approaches to Stop Hypertension) Intervention for Neurodegenerative Delay; PCR, Protein-to-Carbohydrate Ratio.



Supplementary Fig. 12 | Correlation between diet quality-associated lipidomic profiles and CVD mortality risk. Scatter plots show correlations between lipid–diet quality association coefficients and lipid–CVD mortality risk in the AusDiab cohort ($n = 264$ CVD deaths and 8,000 controls over 10 years of follow-up). Each data point represents a lipid species, with the standardized beta coefficient for its association with a diet quality score (x-axis) plotted against the hazard ratio (HR) per 1-SD for CVD mortality derived from univariate Cox proportional hazards models (y-axis). Both lipid–diet and lipid–CVD associations were adjusted for age, sex, BMI, clinical lipids (total cholesterol, HDL-C, and triglycerides), smoking, physical activity, hypertension, lipid-lowering medication use, energy intake, and diabetes status. Pearson correlation coefficients and two-tailed p values are shown in the upper right of each panel.

DGI, Australian Dietary Guideline Index; GDQS, Global Diet Quality Score; HSR, Health Star Rating–Dietary Index; LCD, Low-Carbohydrate Diet; MedDiet, Mediterranean Diet Score; MIND, Mediterranean–DASH (Dietary Approaches to Stop Hypertension) Intervention for Neurodegenerative Delay; PCR, Protein-to-Carbohydrate Ratio.