CVR-2024-1644

Word count
Total word count: 901
Number of references: 6
Number of figures: 1
Figure legend: 250

Rare pathogenic variants in G-protein coupled receptor genes involved in gut-to-host communication are associated with cardiovascular disease risk

Short title: Gut-to-host receptors and cardiovascular disease risk

Leticia Camargo Tavares[1*]; Rikeish R. Muralitharan[1]; Matthew Snelson[1]; Francine Z. Marques[1,2*]

[1]Hypertension Research Laboratory, Victorian Heart Institute and Department of Pharmacology, Biomedical Discovery Institute, Faculty of Medicine, Nursing and Health Sciences, Monash University, Melbourne, Australia; [2]Heart Failure Research Laboratory, Baker Heart and Diabetes Institute, Melbourne, Australia.

***Corresponding authors**: Dr Leticia Camargo Tavares, PhD, Postdoctoral Research Fellow, Hypertension Research Laboratory, School of Biological Sciences, Faculty of Science, Monash University, Melbourne, Australia. Email: leticia.camargotavares1@monash.edu; Prof Francine Marques, Hypertension Research Laboratory, Victorian Heart Institute, Level 2, Victorian Heart Hospital, 631 Blackburn Road Clayton, VIC 3168 Monash University, Melbourne, Australia, Phone: +61-03-9905 6958. E-mail: francine.marques@monash.edu

Funding Information: F.Z.M. is supported by a Senior Medical Research Fellowship from the Sylvia and Charles Viertel Charitable Foundation, a National Heart Foundation Future Leader

Fellowship (105663), and National Health & Medical Research Council (NHMRC) Emerging Leader Fellowship (GNT2017382). M.S. is supported by a National Heart Foundation Postdoctoral Fellowship (106698).

Conflict of interest

None.

Data availability

This research has been conducted using the UK Biobank Resource under Application Number 86879. This work uses data provided by patients and collected by the NHS as part of their care and support. Full methods can be accessed at medRxiv: https://doi.org/10.1101/2024.10.01.24314734.

Acknowledgement

We would like to acknowledge Monash Bioinformatics Platform for access to the M3 servers, as well as the participants of the UK Biobank. This research has been conducted using the UK Biobank Resource under Application Number 86879.

Author contributions

LCT led and FZM supported study design and conception; LCT led the formal data analysis and visualization; LCT led and RRM, MS, and FZM supported data interpretation; LCT led and RRM, MS, and FZM supported drafting the manuscript; all authors approved the final version of the manuscript.

Keywords: Genomics; Rare variants; G-protein couple receptors; Short-chain fatty acids; Hypertension; Major adverse cardiac events.

Gut microbial metabolites called short-chain fatty acids (SCFAs) confer protection against cardiovascular disease (CVD) and its primary risk factor, hypertension.[1] Dietary fibers reach the large intestine undigested, where some are metabolized by fermentative microbes, releasing SCFAs as by-products. These metabolites lower blood pressure in humans and protect against cardiac hypertrophy and fibrosis in mice.[1, 2] Proposed mechanisms include anti-inflammatory signaling activated by the binding of SCFAs to G-protein-coupled receptors (GPCRs, particularly *FFAR3*/GPR41, *FFAR2*/GPR43, and *HCAR2*/GPR109A), as suggested by previous mouse studies using knockout strains.[1, 3] Here, we conducted a large-scale population-based study aimed at determining if rare pathogenic variants (RPVs) affecting SCFA-sensing GPCR genes increase the risk in humans of hypertension and major adverse cardiac events (MACE), including acute coronary syndrome, heart failure, and ischemic stroke.

We used the population healthcare, lifestyle, and whole-exome sequencing (WES) data from the UK Biobank, available for 393,649 participants of white-European genetic ancestry (**Figure A**). UK Biobank has approval from the Northwest Multi-Centre Research Ethics Committee and conforms to the Declaration of Helsinki principles. Informed written consent was given before the inclusion of subjects in the study. Rare (minor allele frequency <1%) variants in the GPCR genes *FFAR3*, *FFAR2* and *HCAR2* and their genotypes were extracted from quality-controlled WES data, based on genomic coordinates. We identified a subset of 158 RPVs with predicted high-impact functional consequences or missense-pathogenic likelihood scores (**Figure A**). Annotations were based on the Ensembl Variant Effect Predictor tool (http://www.ensembl.org/info/docs/tools/vep), as performed previously.[4] Cases and controls for hypertension and MACE subtypes were identified based on diagnoses from hospital admissions (recorded as ICD-10 – International Classification of Disease 10th version - codes), self-reported

medical histories from health-related questionnaires, surgical records and death registries (**Figure A**). The characteristics of the study cohort are detailed in **Figure B**.

Multiple and single GPCR gene-based collapsing analyses were conducted to test whether there is a significant difference in the counts of individuals carrying at least one rare-pathogenic variant that disrupts gene function (carriers) between cases versus controls , using multivariable-adjusted regression models.[5] Age- and sex-adjusted models, either excluding or including BMI as a covariate (*model 1* and *model 2*, respectively), revealed the prevalence of RPV carriers affecting any GPCR gene of interest was significantly higher in participants with hypertension and MACE compared to controls (**Figure C**). Given the occurrence of MACE is recognized to be associated with standard modifiable cardiovascular risk factors (SMuRFs) – including hypertension, diabetes, hypercholesterolemia, and smoking – we further adjusted MACE analyses for these additional covariates (*model 3*). SMuRFs-adjusted analyses confirmed a significantly increased risk of MACE among GPCR RPV carriers, independent of these additional risk factors (**Figure C**). Finally, when GPCR genes were tested individually, carrying RPVs in the *HCAR2* (GPR109A) gene was associated with a markedly increased risk of heart failure, while carrying RPVs in the *FFAR2* (GPR43) gene was associated with an increased risk of hypertension, with associations remaining significant after correction for multiple comparisons (accounting for 3 GPCR genes and distinct models) (**Figure C**).

Next, to identify other diseases and conditions that are more common among hypertension and MACE patients carrying GPCR RPVs (compared to non-carriers) and could contribute to this association, we performed a comorbidity analysis by retrieving hospital inpatient records for 1,259 ICD-10 codes (Chapters I–XIV) across 39,563,559 data-points (243 data-fields covering 162,813 cases for hypertension/MACE). After filtering for ICD-10 diseases

with a prevalence greater than 1% and adjusting for multiple comparisons (297 ICD-10 diseases and two models), we found cardiomyopathy (ICD-10: I42) was more than twice as common in hypertensive patients carrying GPCR RPVs (**Figure D**). Consistently, cardiac hypertrophy and fibrosis have been observed in experimental mouse models that lack SCFA-sensing GPCR genes.[1, 3]

Since SCFAs are produced primarily by the gut microbiota as a response to dietary fiber intake, to further understand the complex relationship between fiber, gut-to-host communication, and cardiovascular health, we investigated whether carrying GPCR RPVs increases the risk of hypertension and MACE among individuals consuming an adequate-fiber diet. This would support the hypothesis that fiber's beneficial cardiovascular effects depend upon GPCR signaling. To estimate daily fiber intake, we studied a subset of UK Biobank participants with dietary data, assessed by a 24-hour dietary recall of food intake on the previous day in five different instances (n=168,142 with dietary data available for at least one instance). Based on nutritional recommendations,[6] an adequate-fiber diet was defined as an average intake of ≥25 g/day for women and ≥30 g/day for men, leaving 13,858 participants for subsequent analyses. Consistent with our hypothesis, the prevalence of hypertension was significantly higher among GPCR RPVs carriers (**Figure E**), suggesting the cardiovascular benefits of fiber intake are diminished in the presence of genetic impairments in SCFA-related GPCR signaling. No significant associations were observed for MACE, likely due to the smaller sample size. Furthermore, in the inadequate-fiber consumers subgroup (n=154,284), hypertension prevalence was still higher in carriers versus non-carriers of RPVs (38.0% versus 34.9%), but it was no longer significant ($p>0.05$). In the absence of sufficient fiber, SCFA production is minimal and,

thus, the benefit from SCFA signaling is likely blunted for both GPCR RPVs carriers and non-carriers.

Overall, our findings elucidate gene-by-environment interactions and show that lack of signaling via SCFA-sensing GPCRs, even when fiber intake is adequate, is associated with human hypertension and MACE in a large population-based cohort from the UK. This underscores the potential of targeting gut-to-host mechanisms through SCFA-sensing GPCRs as a therapeutic strategy for hypertension and CVD.

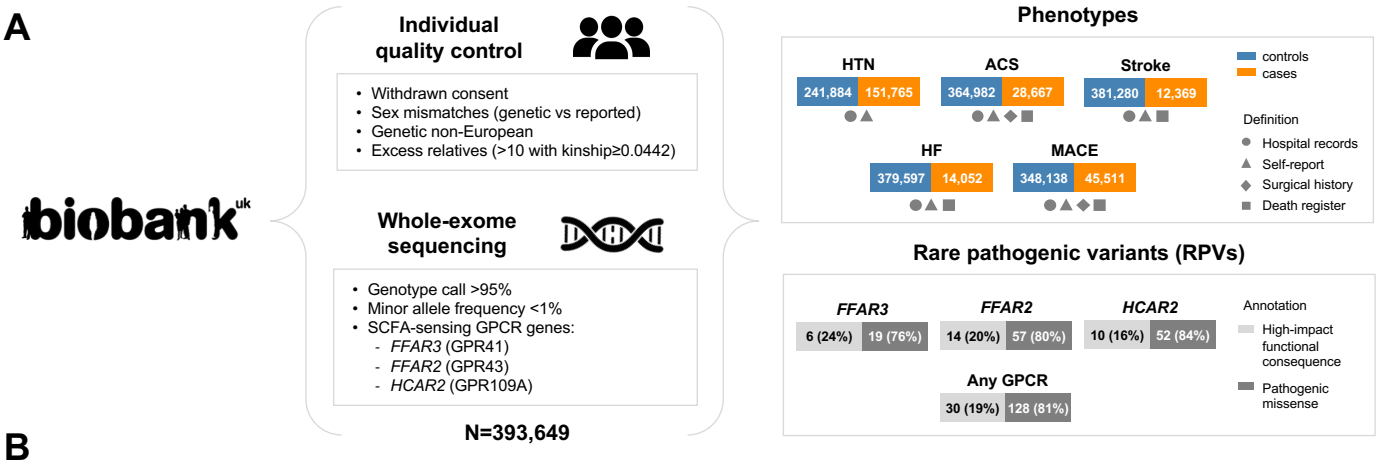
References

1. Kaye DM, Shihata WA, Jama HA, Tsyganov K, Ziemann M, Kiriazis H, Horlock D, Vijay A, Giam B, Vinh A, Johnson C, Fiedler A, Donner D, Snelson M, Coughlan MT, Phillips S, Du XJ, El-Osta A, Drummond G, Lambert GW, Spector TD, Valdes AM, Mackay CR, Marques FZ. Deficiency of Prebiotic Fiber and Insufficient Signaling Through Gut Metabolite-Sensing Receptors Leads to Cardiovascular Disease. *Circulation* 2020;**141**:1393-1403.
2. Jama HA, Rhys-Jones D, Nakai M, Yao CK, Climie RE, Sata Y, Anderson D, Creek DJ, Head GA, Kaye DM, Mackay CR, Muir J, Marques FZ. Prebiotic intervention with HAMSAB in untreated essential hypertensive patients assessed in a phase II randomized trial. *Nature Cardiovascular Research* 2023;**2**:35-43.
3. R Muralitharan R, Zheng T, Dinakis E, Xie L, Barbaro-Wahl A, Jama HA, Nakai M, Patterson M, Leung KC, McArdle Z, Mirabito Colafella K, Johnson C, Qin W, Salimova E, Bitto N, Liaskos MK, Kaye DM, O'Donnell JA, Mackay CR, Marques FZ. Gut Microbiota Metabolites Sensed by Host GPR41/43 Protect Against Hypertension. *Circulation Research* 2025.
4. Zheng T, Camargo-Tavares L, Bonfiglio F, Marques FZ, Naim HY, D'Amato M. Rare Hypomorphic Sucrase Isomaltase Variants in Relation to Irritable Bowel Syndrome Risk in UK Biobank. *Gastroenterology* 2021;**161**:1712-1714.
5. Povysil G, Petrovski S, Hostyk J, Aggarwal V, Allen AS, Goldstein DB. Rare-variant collapsing analyses for complex traits: guidelines and applications. *Nature Reviews Genetics* 2019;**20**:747-759.

6. Jama HA, Snelson M, Schutte AE, Muir J, Marques FZ. Recommendations for the Use of Dietary Fiber to Improve Blood Pressure Control. *Hypertension* 2024;**81**:1450-1459.

Figure legend

Figure. Rare pathogenic variants (RPVs) in short-chain fatty acid (SCFA)-sensing G-protein coupled receptors (GPCR) genes are associated with poor cardiovascular outcomes. **(A)** Study design flowchart, including phenotype and RPVs definitions. RPVs were identified based on two classifications: (1) *'high-impact functional consequence'*, comprising stop gained/lost, frameshift, and splicing alterations based on Ensembl Variant Effect Predictor annotations; (2) *'pathogenic missense variants'*, comprising missense variants with pathogenic likelihood scores >0.564 from AlphaMissense, >0.025 from Mendelian Clinically Applicable Pathogenicity, or >20 from Combined Annotation Dependent Depletion. **(B)** Study cohort demographic and clinical characteristics, stratified as cases and controls for each phenotype of interest, encompassing hypertension (HTN), major adverse cardiac events (MACE) and MACE subtypes (acute coronary syndrome (ACS), heart failure (HF), and ischemic stroke). **(C)** Significant associations between GPCR RPVs and phenotypes' risk. **(D)** Increased cardiomyopathy prevalence in hypertensive cases carrying GPCR RPVs. **(E)** Increased hypertension prevalence in GPCR RPV carriers among adequate-fiber diet consumers. **Statistics:** Multiple comparisons were corrected by false discovery rate using the Benjamini-Hochberg method (threshold q<0.05). **Panels C-E:** Odds-ratio (OR) and P values are derived from multivariable logistic regression analyses, adjusted for: age and sex (*model 1*); age, sex, and BMI (*model 2*); age, sex, BMI, and standard modifiable cardiovascular risk factors (SMurFs), including smoking (current/past), hypertension, diabetes, and hypercholesterolemia (*model 3*; only applied for MACE and subtypes). Genetic ancestry (10-top genetic principal components) was also included as covariate when testing the association between GPCR RPVs and phenotypes' risk (Panel C). All OR forest plots are in a $\log_{10}$ scale.

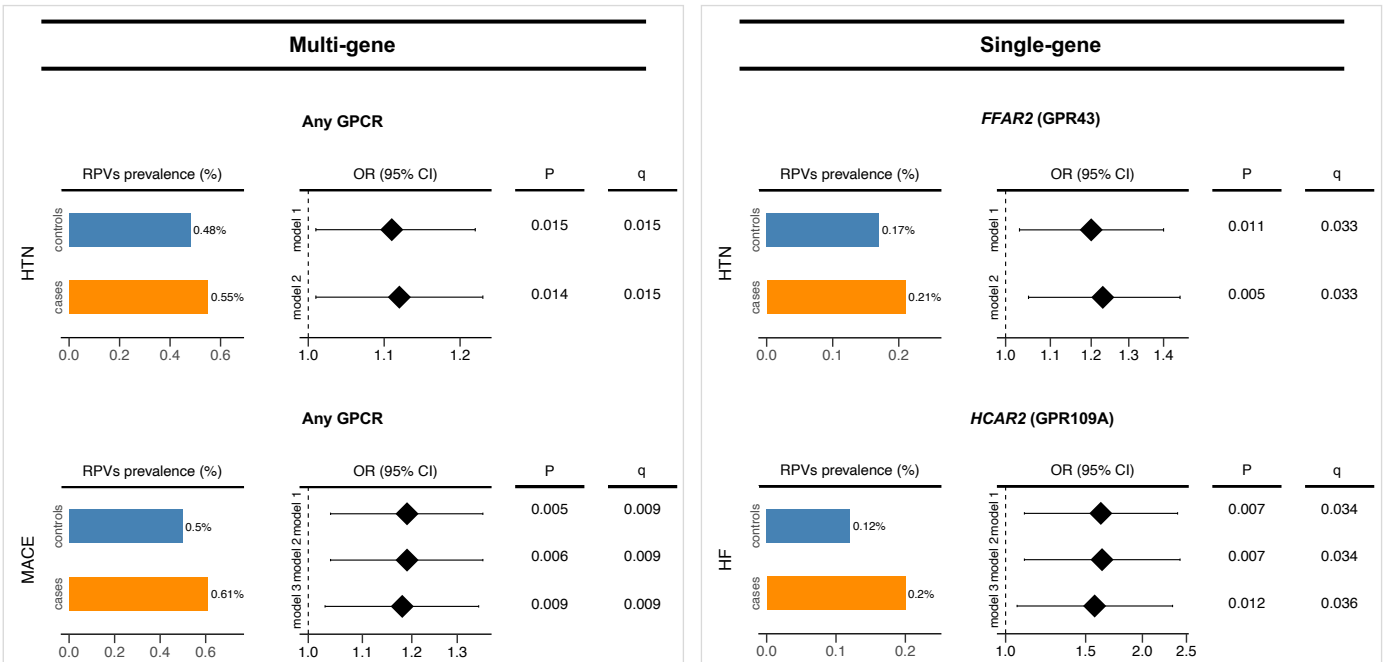


B

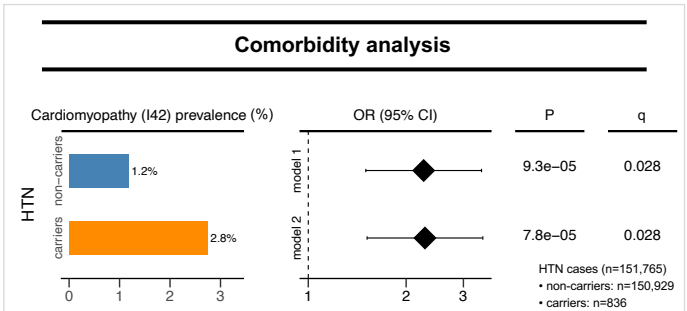
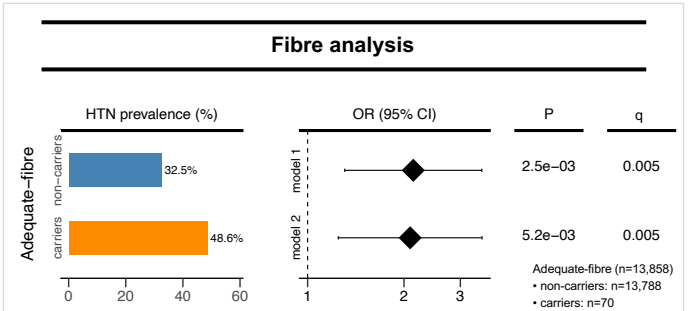
	HTN		ACS		Stroke		HF		MACE	
	controls	cases	controls	cases	controls	cases	controls	cases	controls	cases
Demographical and clinical data										
Women (n, %)	141,134 (58.3)	71,584 (47.2)	205,130 (56.2)	7,588 (26.5)	207,841 (54.5)	48,77 (39.4)	207,803 (54.7)	4,915 (35.0)	197,861 (56.8)	14,857 (32.6)
Age, years (mean, SD)	55.07 (8.06)	59.85 (6.96)	56.58 (8.02)	61.18 (6.43)	56.77 (8.00)	61.48 (6.50)	56.72 (8.00)	62.30 (6.00)	56.34 (8.01)	61.30 (6.45)
Body mass index (BMI), kg/m^2 (mean, SD)	26.39 (4.21)	29.05 (5.08)	27.30 (4.72)	28.90 (4.75)	27.38 (4.73)	28.61 (5.00)	27.33 (4.68)	29.82 (5.74)	27.22 (4.67)	28.92 (5.03)
Systolic blood pressure, mmHg (mean, SD)	134.33 (17.36)	149.48 (19.42)	139.90 (19.54)	144.10 (20.23)	140.04 (19.56)	145.19 (20.73)	140.02 (19.53)	145.14 (21.37)	139.63 (19.45)	144.60 (20.40)
Diastolic blood pressure, mmHg (mean, SD)	79.83 (9.76)	86.09 (10.85)	82.29 (10.60)	81.80 (11.19)	82.22 (10.62)	83.18 (11.26)	82.26 (10.60)	82.07 (11.66)	82.23 (10.57)	82.40 (11.22)
Smoking (current or past) (n, %)	69,997 (28.9)	58,082 (38.3)	113,954 (31.2)	14,125 (49.3)	122,407 (32.1)	5,672 (45.9)	121,062 (31.9)	7,017 (49.9)	106,536 (30.6)	21,543 (47.3)
Hypertension (n, %)	0 (0.0)	151,765 (100.0)	129,456 (35.5)	22,309 (77.8)	142,578 (37.4)	9,187 (74.3)	140,422 (37.0)	11,343 (80.7)	117,302 (33.7)	34,463 (75.7)
Diabetes (n, %)	7,622 (3.2)	26,994 (17.8)	27,695 (7.6)	6,921 (24.1)	31,806 (8.3)	2,810 (22.7)	30,330 (8.0)	4,286 (30.5)	24,115 (6.9)	10,501 (23.1)
Hypercholesterolemia (n, %)	20,875 (8.6)	59,971 (39.5)	60,530 (16.6)	20,316 (70.9)	74,002 (19.4)	6,844 (55.3)	72,775 (19.2)	8,071 (57.4)	52,896 (15.2)	27,950 (61.4)
Any SMuRF (n, %)	85,398 (35.3)	151,765 (100.0)	210,340 (57.6)	26,823 (93.6)	226,078 (59.3)	11,085 (89.6)	224,160 (59.1)	13,003 (92.5)	195,517 (56.2)	41,646 (91.5)
GPCR RPVs carriers										
Any GPCR (GPR 41/43/109A) (n, %)	1167 (0.48)	836 (0.55)	1833 (0.50)	170 (0.59)	1932 (0.51)	71 (0.57)	1914 (0.50)	89 (0.63)	1725 (0.50)	278 (0.61)
FFAR3 (GPR41) (n, %)	490 (0.20)	319 (0.21)	747 (0.20)	62 (0.22)	782 (0.21)	27 (0.22)	775 (0.20)	34 (0.24)	703 (0.20)	106 (0.23)
FFAR2 (GPR43) (n, %)	413 (0.17)	315 (0.21)	662 (0.18)	66 (0.23)	700 (0.18)	28 (0.23)	701 (0.18)	27 (0.19)	626 (0.18)	102 (0.22)
HCAR2 (GPR109A) (n, %)	269 (0.11)	204 (0.13)	431 (0.12)	42 (0.15)	457 (0.12)	16 (0.13)	445 (0.12)	28 (0.20)	403 (0.12)	70 (0.15)

Quantitative variables are shown as 'mean (SD)' and qualitative as 'n (%)'

C



D



E