



OPEN

Evidence for involvement of the alcohol consumption *WDPCP* gene in lipid metabolism, and liver cirrhosis

Felix O'Farrell^{1,13}, Benjamin Aleyakpo^{2,13}, Rima Mustafa^{3,4,13}, Xiyun Jiang^{1,13}, Rui Climaco Pinto^{3,5,13}, Paul Elliott^{3,4,5,6,7,8}, Ioanna Tzoulaki^{3,9}, Abbas Dehghan^{3,4,5}, Samantha H. Y. Loh¹⁰, Jeff W. Barclay¹¹, L. Miguel Martins¹⁰ & Raha Pazoki^{1,3,12}✉

Biological pathways between alcohol consumption and alcohol liver disease (ALD) are not fully understood. We selected genes with known effect on (1) alcohol consumption, (2) liver function, and (3) gene expression. Expression of the orthologs of these genes in *Caenorhabditis elegans* and *Drosophila melanogaster* was suppressed using mutations and/or RNA interference (RNAi). In humans, association analysis, pathway analysis, and Mendelian randomization analysis were performed to identify metabolic changes due to alcohol consumption. In *C. elegans*, we found a reduction in locomotion rate after exposure to ethanol for RNAi knockdown of *ACTR1B* and *MAPT*. In *Drosophila*, we observed (1) a change in sedative effect of ethanol for RNAi knockdown of *WDPCP*, *TENM2*, *GPN1*, *ARPC1B*, and *SCN8A*, (2) a reduction in ethanol consumption for RNAi knockdown of *TENM2*, (3) a reduction in triacylglycerols (TAG) levels for RNAi knockdown of *WDPCP*, *TENM2*, and *GPN1*. In human, we observed (1) a link between alcohol consumption and several metabolites including TAG, (2) an enrichment of the candidate (alcohol-associated) metabolites within the linoleic acid (LNA) and alpha-linolenic acid (ALA) metabolism pathways, (3) a causal link between gene expression of *WDPCP* to liver fibrosis and liver cirrhosis. Our results imply that *WDPCP* might be involved in ALD.

Alcohol consumption is a major public health concern and is responsible for over 5% of the global burden of disease¹. It has been known for a long time that excessive drinking leads to a range of molecular changes including acceleration of hepatic lipogenesis which leads to liver pathologies, such as liver fibrosis and liver cirrhosis commonly known as alcoholic liver disease (ALD). Although some of the known alcohol behavior genes such as *ADH* and *CYP2E1* genes have a known role in hepatic lipogenesis², the full picture of the biological pathways and molecular changes occurring as a result of alcohol consumption that leads to hepatic lipogenesis is not fully understood.

¹Cardiovascular and Metabolic Research Group, Division of Biosciences, Department of Life Sciences, College of Health and Life Sciences, Brunel University, London UB8 3PH, UK. ²The Francis Crick Institute, London NW1 1AT, UK. ³Department of Epidemiology and Biostatistics, School of Public Health, St Mary's Campus, Imperial College London, Norfolk Place, London W2 1PG, UK. ⁴UK Dementia Research Institute, Imperial College London, Exhibition Road, London SW7 2AZ, UK. ⁵MRC Centre for Environment and Health, Department of Epidemiology and Biostatistics, School of Public Health, St Mary's Campus, Imperial College London, Norfolk Place, London W2 1PG, UK. ⁶British Heart Foundation Centre of Research Excellence, Imperial College London, Du Cane Road W12 0NN, UK. ⁷National Institute for Health Research, Imperial Biomedical Research Centre, Imperial College London, Exhibition Road, London SW7 2AZ, UK. ⁸Health Data Research UK at Imperial College London, Exhibition Road, London SW7 2AZ, UK. ⁹Centre for Systems Biology, Biomedical Research Foundation, Academy of Athens, Athens, Greece. ¹⁰MRC Toxicology Unit, University of Cambridge, Gleeson Building, Tennis Court Road, Cambridge CB2 1QR, UK. ¹¹Department of Molecular Physiology and Cell Signalling, Institute of Systems, Molecular and Integrative Biology, University of Liverpool, Liverpool L69 3BX, UK. ¹²Division of Biomedical Sciences, Department of Life Sciences, College of Health and Life Sciences, Brunel University, London UB8 3PH, UK. ¹³These authors contributed equally: Felix O'Farrell, Benjamin Aleyakpo, Rima Mustafa, Xiyun Jiang and Rui Climaco Pinto. ✉email: raha.pazoki@brunel.ac.uk

Advances in omics such as genomics and metabolomics within the last two decades have resulted in a boost in our understanding of the mechanism of diseases through agnostic approaches such as genome-wide association studies (GWAS) that revealed numerous genetic loci linked to complex diseases. Recently, we have applied GWAS to identify genetic variants in the form of single nucleotide polymorphisms (SNPs) that are associated with alcohol consumption³ as well as with circulating liver enzymes⁴ in the European populations. Some of the identified alcohol genes (e.g., *ADH*, *KLB*, *DRD2*) have been investigated for a better understanding of their involvement in alcohol consumption and health consequences such as hepatic lipogenesis. However, the biological effect of most of the identified alcohol-associated genes remains to be elucidated.

In this study, we aimed to shed light on the biological pathways and molecular changes linking alcohol consumption and liver pathologies. We first investigated molecular consequences of alcohol consumption and the pathways involved at metabolic level using omics approaches. Subsequently in search for common pathways between alcohol consumption and liver pathologies, we identified candidate genes with effect on both alcohol consumption and liver function and then investigated the biological effect of candidate genes in model organism i.e., ethanol-exposed *C. elegans* and *Drosophila* to generate knowledge that could ultimately be used to better understand alcohol related behavior and hepatic lipogenesis. We finally returned to use data from humans to further validate the most plausible candidate genes with evidence of potential involvement in lipogenesis.

Methods

Population

In the current study, we followed a multi-stage approach using population-based studies and model organisms to better understand pathways involved in alcohol consumption and its health consequences (Fig. 1). We used data from the Airwave Health Monitoring Study⁵, an occupational cohort of 53,116 police officers and staff ages 18 years and over across the UK (Supplementary Table 1). The Airwave Health Monitoring Study was approved by the UK National Research Ethics Service (NRES) North West—Haydock Regional Ethics Committee (REC reference: NRES/19/NW/0054; IRAS ID: 259978). The North-West Haydock REC approved the study protocol and all study documentation, and prior individual informed consent was obtained from each of the study participants. All methods were carried out in accordance with relevant guidelines and regulations. Participants were informed about the study and provided informed consent. Detailed information about the Airwave population, metabolic assays, data processing, metabolite annotation as well as genotyping and imputation is included in the “Supplementary Methods”.

Alcohol consumption

Alcohol consumption during the last seven days was assessed through a self-reported questionnaire in the Airwave study. The participants were asked to quantify (1) the number of glasses (small/125 ml) of red wine, white wine/champagne, fortified wine (includes sherry, port and vermouth), (2) the number of pub measures of spirits/liqueurs (includes whisky, gin, rum, vodka and brandy), and (3) the number of pints of beer or cider (include bitter, lager, stout, ale and Guinness) consumed in the last seven days. We considered the density of alcohol to be 0.79 g/ml at room temperature and calculated the total amount of alcohol consumed by summing up the grams of alcohol consumed from each category of alcoholic drink.

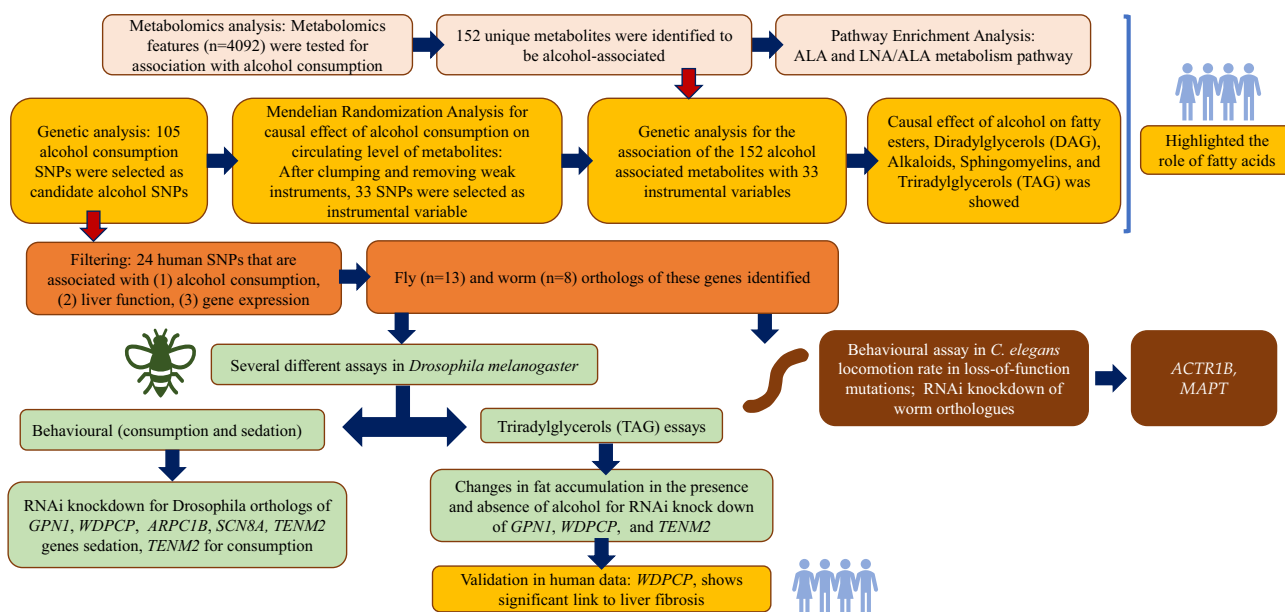


Figure 1. Overview of the study design and the findings.

Analysis of the metabolome

We analyzed the association between alcohol consumption and circulating metabolites within the Airwave sample to identify candidate metabolites associated with alcohol consumption within the Airwave sample (alcohol-associated metabolites). Metabolomics features ($n = 4092$) were obtained from data acquired by the National Phenome Centre (NPC) using liquid-chromatography/mass spectrometry, covering a wide range of hydrophilic and lipid metabolite classes. We excluded unannotated features leaving 986 annotated metabolites for the analysis and tested the association of these metabolites with alcohol consumption. Subsequently, to assess if the resulting alcohol-associated metabolites were additionally involved in known metabolic pathways, we performed pathway enrichment analysis using all annotated alcohol-associated metabolomics features. To this end, Kyoto Encyclopedia of Genes and Genomes (KEGG)⁶ and The Small Molecule Pathway Database (SMPDB)⁷ were used and the statistical significance level was claimed using a P -value calculated using false discovery rate (P_{dr}).

Genotyping and imputation

In the Airwave study⁵, genotyping was conducted using Illumina Infinium HumanExome-12v1-1 BeadChip Array. QC steps were performed to remove any samples with high missingness ($> 3\%$) or outlier heterozygosity rates ($> 3SD$ from the batch mean), duplicates (ID-based or genotype-based) or presented with high degree of relatedness. Markers with genotype call rate $< 98\%$, significant deviation from Hardy–Weinberg equilibrium ($P < 1 \times 10^{-5}$), and low MAF ($< 1\%$) were removed. Genotype data was imputed with 1000 Genomes phase 3 reference panel.

GWAS on metabolomics

We performed genome-wide association studies (GWAS) in the Airwave Study to obtain genetic association estimates for alcohol-related SNPs and selected metabolomic features. Untargeted Mass-Spectrometry (MS) was performed in heparin plasma samples of the Airwave Study, producing three separate datasets, namely hydrophilic interaction liquid chromatography (HILIC) positive (HPOS), lipid positive (LPOS), and lipid negative (LNEG). For each metabolomic dataset, we used principal component analysis (PCA) to identify outlier samples and exclude them. The data was residualized using the first ten principal components to account for population stratification. We transformed the data into z -scores, performed individually for each metabolomic feature, using median and median absolute deviations (MAD). Values that were more than 5 MAD from the median were removed and imputed using k -nearest neighbors' imputation⁸. GWAS were conducted for each metabolomic feature with adjustment for age and sex ($N = 1970$). Due to the high number of metabolomic features, GWAS were performed using the high-dimensional association analyses (HASE) framework, which applies matrix operation and removes redundant calculation in high-dimensional association analyses to improve computational efficiency⁹.

We used the results from association analyses between the alcohol-associated metabolites and the known alcohol-associated SNPs to further perform a causal inference analysis on these results using the inverse variance weighted two-sample Mendelian randomization (MR) analysis¹⁰. The aim of the MR analysis was to identify changes in circulating metabolites caused by alcohol consumption. MR is a causal assessment method used in observational studies to mimic randomized controlled trials (RCTs) by taking advantage of the random assortment of alleles at conception. MR uses instrumental variables (i.e., genotype status), that are robustly associated with an exposure of interest, as a natural randomization tool occurring at conception¹¹. The detailed methods for selection of alcohol consumption instruments have been described previously^{10,12}. Briefly, we obtained genetic association statistics (β values) for 105 alcohol-associated SNPs^{3,13} from Liu and colleagues¹³. Clumping and removing weak instruments as described previously^{10,12} ensured the most robust instruments were used for the MR analysis ($n = 33$).

Gene selection for model organisms

We selected a list of 105 alcohol-related SNPs from recently conducted GWAS of alcohol consumption^{3,13}. SNPs were selected if they presented a P -value lower than a GWAS significance threshold of 5×10^{-8} in their association with alcohol consumption. As we were interested in finding pathways involved in hepatic lipogenesis as a result of alcohol consumption, we sought candidate SNPs that showed strong effect with both alcohol consumption and liver function, using our recently published GWAS of circulating liver enzymes⁴. To account for multiple testing, a corrected P -value threshold of < 0.00048 was used for the association with liver enzymes. This P -value threshold corresponds to a nominal P -value (0.05) that has been adjusted for the number of alcohol-related SNPs ($n = 105$) using the Bonferroni method¹⁴. SNPs ($n = 43$; Supplementary Table 2) that were associated with at least one of the three liver enzymes alanine transaminase (ALT), alkaline phosphatase (ALP), and gamma-glutamyl transferase (GGT) were shortlisted to assess their link to gene expression using the GTEx database¹⁵. We eventually selected SNPs ($n = 24$) that demonstrated evidence of a statistically significant effect on gene expression of their nearest genes (a cis-eQTL effect; Table 1).

Ortholog selection

The 24 SNPs with cis-eQTL effect with their annotated genes were matched to their orthologs in *Drosophila* using FlyBase (DIOPT online tool version 8.5/9.0; beta; <http://www.flyrnai.org/diopt>). To ensure the selection of the most credible orthologs, we used scores calculated in FlyBase. This database provides a number of approaches that support the gene-pair relationship out of a total number of tools that computed relationships between *Homo sapiens* and *Drosophila*. Genes with a score of > 12 were shortlisted for further analysis in *Drosophila*. Nineteen *Drosophila* orthologs were identified of which eight had a score ≥ 12 including *ARPC1B* (*arpc1*), *ACTR1B* (*arp1*), *GPN1* (*CG3704*), *WDPCP* (*frz*), *MLXIPL* (*mondo*), *SLC4A8* (*ndae1*), *SCN8A* (*para*), and *TENM2* (*Ten-m*).

Alcohol SNP *	Annotated gene	eQTL tissue	eQTL effect size§ (effect allele)	eQTL P value†	Drosophila ortholog‡	FlyBase score out of 15§	C. elegans ortholog (score)‡	Drosophila strains	C. elegans strain
rs823114	<i>NUCKS1</i>	Thyroid	-0.16 (A)	5.80×10^{-18}	-	-	-	-	-
rs1260326	<i>GCKR</i>	Thyroid	0.22 (C)	1.80×10^{-08}	-	-	-	-	-
rs2178197	<i>GPN1</i>	Brain-Cerebellar Hemisphere	-0.31 (G)	6.50×10^{-08}	<i>cg3704</i>	13	<i>gop-2</i> (9) †	CG3704 (55,294)	RG5036 <i>gop-2</i> (<i>gk5528</i>),
rs13032049	<i>WDPCP</i>	Adipose-Subcutaneous	-0.16 (G)	3.00×10^{-04}	<i>fritz</i>	13	-	Frz (55,649)	-
rs11692435	<i>ACTR1B</i>	Thyroid	0.89 (A)	1.50×10^{-80}	<i>arp1</i>	12	<i>arp-1</i> (8) †	Lethal, not studied	FX4735 <i>arp-1</i> (<i>tm4735</i>)
rs12646808	<i>MSANTD1</i>	Thyroid	-0.25 (T)	1.30×10^{-09}	<i>cg18766</i>	8	-	-	-
rs11940694	<i>KLB</i>	Liver	0.22 (A)	4.00×10^{-05}	<i>cg9701</i>	6	<i>klo-1</i> (5), <i>klo-2</i> (5)	-	Not studied
rs1229984	<i>ADH1B</i>	Esophagus-Gastroesophageal Junction	-1.4 (C)	1.90×10^{-16}	<i>fdh</i>	7	<i>adh-5</i> (5)	-	Not studied
rs10078588	<i>TENM2</i>	Thyroid	0.29 (A)	2.80×10^{-15}	<i>ten-m</i>	13	<i>ten-1</i> (10) †	<i>TenM</i> (29,390)	VC518 <i>ten-1</i> (<i>ok641</i>),
rs34060476	<i>MLXIPL</i>	Pancreas	0.51 (G)	4.50×10^{-18}	<i>mondo</i>	12	<i>mml-1</i> (9)	<i>Mondo</i> (27,059)	RB954 <i>mml-1</i> (<i>ok849</i>)
rs10249167	<i>ARPC1B</i>	Thyroid	0.3 (G)	1.80×10^{-25}	<i>arpc1</i>	13	<i>arx-3</i> (10) †	<i>Arpc1</i> (31,246)	VC3166 <i>arx-3</i> (<i>ok1122</i>)
rs988748	<i>BDNF</i>	Brain-Nucleus accumbens basal ganglia	0.22 (G)	9.70×10^{-06}	-	-	-	-	-
rs2071305	<i>MYBPC3</i>	Whole blood	-0.13 (C)	2.30×10^{-11}	<i>hbs</i>	1	-	-	-
rs7121986	<i>DRD2</i>	Esophagus-Muscularis	0.26 (C)	3.50×10^{-08}	<i>dop2r</i>	9	<i>dop-2</i> (6), <i>dop-3</i> (6)	-	Not studied
rs10876188	<i>SLC4A8</i>	Cells-Cultured fibroblasts	-0.12 (T)	1.80×10^{-07}	<i>ndae1</i>	12	<i>abts-1</i> (9)	Lethal, not studied	RB1381 <i>abts-1</i> (<i>ok1566</i>)
rs7958704	<i>SCN8A</i>	Nerve-Tibial	-0.25 (C)	2.90×10^{-10}	<i>para</i>	12	<i>cca-1</i> (1), <i>unc-77</i> (1), <i>egl-19</i> (4)	<i>Para</i> (33,923)	JD21 <i>cca-1</i> (<i>ad1650</i>)
rs12312693	<i>STAT6</i>	Brain-Frontal Cortex BA9	0.23 (C)	8.00×10^{-05}	<i>stat92e</i>	9	<i>sta-2</i> (4), <i>sta-1</i> (7)	-	RB796 <i>sta-1</i> (<i>ok587</i>)
rs11625650	<i>KIF26A</i>	Whole blood	-0.26 (A)	9.10×10^{-12}	<i>cg14535</i>	7	<i>vab-8</i> (2)	-	NG2484 <i>vab-8</i> (<i>gm84</i>)
rs1421085	<i>FTO</i>	Muscle-Skeletal	0.14 (C)	7.50×10^{-08}	-	-	-	-	-
rs11648570	<i>PMFBP1</i>	Esophagus-Mucosa	-0.29 (C)	9.00×10^{-05}	<i>cg12702</i>	1	-	-	-
rs3803800	<i>TNFSF13</i>	Brain-Cortex	-0.2 (G)	2.50×10^{-08}	<i>egr</i>	3	-	-	-
rs1053651	<i>TCAP</i>	Muscle-Skeletal	-0.11 (C)	4.90×10^{-11}	-	-	-	-	-
rs1991556	<i>MAPT</i>	Brain-Cerebellum	-0.28 (A)	1.80×10^{-06}	<i>tau</i>	7	<i>ptl-1</i> (8)	-	RB809 <i>ptl-1</i> (<i>ok621</i>)
rs4815364	<i>ACSS1</i>	Brain-Cerebellar Hemisphere	-0.47 (A)	1.70×10^{-12}	<i>accoas</i>	6	-	-	-

Table 1. Overview of the genetic variants with effect on alcohol consumption, liver enzymes and gene expression. eQTL expression quantitative trait loci. *Alcohol SNPs associated with liver enzymes at replication P value < 0.000476. Liver enzyme GWAS summary statistics obtained from Pazoki et al. 2021. †Normalized effect size and P-value obtained from GTEx portal. ‡Orthologs were extracted from FlyBase and WormBase. § number of FlyBase tools that support the gene-pair relationship out of a total number of tools (n = 15) that computed relationships between *Homo sapiens* and *Drosophila melanogaster*. †Genes for which mutations were homozygous lethal and as such, heterozygote mutations balanced by chromosomal translocations were instead analyzed.

To identify potential worm orthologs, the SNPs with cis-eQTL effect with their annotated genes were sought for their worm orthologs within WormBase (<https://wormbase.org/>). BLASTp analysis results of human protein sequence against *C. elegans* protein database were obtained in WormBase version WS280 using data from *C. elegans* Sequencing Consortium genome project (PRJNA13758). The human protein sequence for the protein encoded by the genes under this study was extracted from UniProt (<https://www.uniprot.org/>). Worm genes with the best alignment with the human protein sequence indicated by an Expect value (E value) < 1×10^{-5} and the highest BLASTp score (bits) were moved forward. The E value represents the number of alignments that could be found in similarity to the protein sequence by chance. Using WormBase, 13 worm orthologs were identified including *ACTR1B* (*arp-1*), *ARPC1B* (*arx-3*), *GPN1* (*gop-2*), *MLXIPL* (*mml-1*), *STAT6* (*sta-1*), *TENM2* (*ten-1*), *KIF26A* (*vab-8*), *SLC4A8* (*abts-1*), *MAPT* (*ptl-1*) and *SCN8A* (*cca-1*). We excluded *KLB* (*klo-1*), *DRD2* (*dop-2*), and *ADH1B* (*adh-5*) from the *C. elegans* experiments as similar studies investigating these genes already exist^{16–18}.

We did not identify any fly or worm orthologs for *NUCKS1*, *GCKR*, *BDNF*, *FTO*, *TCAP*. Additionally, no worm ortholog was found for *WDPCP*, *MSANTD1*, *MYBPC3*, *PMFBP1*, *TNFSF13*, and *ACSS1*.

Drosophila

Genetics and Drosophila melanogaster strains

All *Drosophila* stocks and crosses were maintained on standard cornmeal agar media at 25 °C on 12/12 h light/dark cycles. The following strains were used as positive control: *w^{berlin}*, *hppy¹⁷⁻⁵¹*, + (*hppy* mutant) and *w^{berlin}*; *hppy¹⁷⁻⁵¹*; *hppy III* (*hppy* mutant with a rescue genomic construct) (gifts from Prof. Ulrike Heberlein, Janelia Research Campus, Virginia, USA), RNAi lines were obtained from Bloomington *Drosophila* Stock Center (gene CG3704 (BDSC no: 55294), *Frz* (BDSC no: 55649), *Mondo* (BDSC no: 27059), *Arpc1* (BDSC no: 31246), *Para* (BDSC no: 33923) and *TenM* (BDSC no: 29390)). All lines used were backcrossed to *w¹¹¹⁸* or [*v*]^{*w1118*} (RNAi lines). The expression of all RNAi constructs was driven by the ubiquitous driver, *daGal4*, which drives expression throughout development from embryonic to adult stage in all tissues. All the experiments on adult flies were performed using males.

Drosophila ethanol consumption assay

The Capillary FEeder (CAFE) assay¹⁹ was used to measure ethanol consumption. Eight male flies were placed into an experimental vial (8 cm height, 3.3 cm diameter) containing 6 microcapillary tubes (BRAND® disposable BLAUBRAND® micropipettes, intraMark, BR708707, with 1 µl marks), each containing 5 µl of liquid food. Liquid food was prepared by dissolving 50 mg of yeast granules in 1 ml of boiling water by vortexing, followed by brief centrifugation. Then, 40 mg of sucrose (Sigma–Aldrich, 84097) was added to 800 µl of the dissolved yeast mixture, followed by vortexing. The microcapillary tubes were filled with liquid food up to the 5 µl mark. Ethanol food consists of normal food supplemented with 15% ethanol. Each experiment consisted of 5 experimental vials per genotype with each vial containing normal food (3 capillaries) and ethanol food (3 capillaries). The flies were acclimatized in the experimental vial without any food for 2 h prior to the start of the experiment. This step was also used to incentivise the flies to eat once the food was introduced. The experimental vials were placed in a plastic box with a cover to control humidity. The flies were allowed to feed for 19 h, after which the amount consumed (in mm) was measured with a digital calliper (Dasqua Bluetooth Digital Calliper 12"/300 mm, 24108120). The total amount of food consumed was calculated using the formula:

$$\text{Food uptake } (\mu\text{l}) = (\Sigma \text{ measured distances between 3 microcapillary tubes (mm)}) / 14.42 \text{ mm (1 } \mu\text{l in measured distance)} / 8 \text{ flies.}$$

Drosophila ethanol sedation assay

Fly sedation assay was performed as previously described²⁰. Briefly, 8 flies were transferred to a 25 mm x 95 mm transparent plastic vial in between two cotton plugs. A piece of cotton plug at the base of the vial served as a stable surface to observe the flies and another plug was used to cap the vial and deliver the ethanol. 500 µl of 100% ethanol was added to the side of the cotton plug facing the flies. Sedation was observed manually as ST50, which is the time in minutes it takes for 50% of the flies in a sample vial to become sedated. Sedation events are recorded when the flies become inactive and lay on their backs for over 10 s.

Drosophila triradylglycerols (TAG) measurements

To assess the role of the candidate genes in lipogenesis, we assessed the effect of RNAi knockdown of the selected genes on TAG levels in *Drosophila*. Eight male flies of the indicated genotypes were placed into an experimental vial as described in the ethanol consumption assay, with all 6 microcapillary tubes filled with normal food (5% sucrose + 5% yeast) or ethanol food (normal food + 15% ethanol), for 2 days. TAGs were assessed through colorimetric assays using 96-well microtiter plates and an Infinite M200Pro multifunction reader (TECAN). The assays were performed as previously described²¹. Briefly, flies were homogenized in 110 µl of PBS + 0.05% Tween 20 (PBST) for 2 min on ice and immediately incubated at 70 °C for 10 min to inactivate endogenous enzymatic activity. A 35 µl fly homogenate sample and a glycerol standard (Sigma, no. G7793) were incubated together with either 35 µl of PBST (for free glycerol measurements) or 35 µl of TAG reagent (Sigma, no. T2449, for TAG measurements) at 37 °C for 60 min. After 3 min of centrifugation at full speed, 30 µl of each sample was transferred into a clear-bottom plate (two technical replicates per biological sample) together with 100 µl of free glycerol reagent (Sigma–Aldrich, F6428) and incubated at 37 °C for 5 min. TAG absorbance was divided by the protein concentration of the respective sample, which was measured by Bradford assay (Sigma–Aldrich, B6916).

C. elegans

Nematode strains and culture

All *C. elegans* strains were cultured on nematode growth medium (NGM) agar plates at 20 °C using *Escherichia coli* OP50 as a food source. For the wildtype worms, Bristol N2 strain was used. Genes for which mutations were homozygous lethal (*arp-1*, *arx-3*, *gop-2* and *ten-1*), heterozygote mutations balanced by chromosomal translocations were instead analyzed. Loss-of-function mutations were not backcrossed into Bristol N2, but instead initially screened for phenotypic differences with those differences validated by RNAi.

Nematode RNA interference (RNAi)

RNAi experiments were performed on the NL2099 *rrf-3* (*pk1426*) strain as previously described^{16,22} (Supplementary Table 3). RNAi was achieved by feeding²³ using the ORFeome based RNAi library²⁴. In brief, HT115 RNAi bacterial clones were initially cultured in LB media with 100 µg/ml ampicillin and subsequently spotted in three 50 µl drops on 60 mm diameter NGM plates containing 1 mM isopropyl β-1-thiogalactopyranoside (IPTG) and 25 µg/ml carbenicillin. Plates were left to dry for 4–7 days before seeding to improve RNAi efficiency. Following seeding, five L3–L4 worms were added to each RNAi plate and cultured at 20 °C until the F1 generation reached adulthood. Ethanol experiments were performed and analyzed as described above and compared to worms fed with an empty RNAi feeding vector.

Nematode behavioural assays

All ethanol experiments were performed at 20 °C in a temperature-controlled room as previously described^{16,22}. Behavioural assays were conducted on young adult hermaphrodites selected from sparsely populated NGM plates. Nematodes with *loss-of-function* mutations in worm orthologues of *ACTR1B* (*arp-1*), *ARPC1B* (*arx-3*), *GPN1* (*gop-2*), *MLXIPL* (*mml-1*), *STAT6* (*sta-1*), *TENM2* (*ten-1*), *KIF26A* (*vab-8*), *SLC4A8* (*abts-1*), *MAPT* (*ptl-1*) and *SCN8A* (*cca-1*) were acutely exposed to ethanol and the resultant effect on rate of locomotion (thrashes per minute) was quantified in Dent's solution (140 mM NaCl, 6 mM KCl, 1 mM CaCl₂, 1 mM MgCl₂, 5 mM HEPES, pH 7.4 with bovine serum albumin at 0.1 mg/ml) by measuring thrashes per minute (one thrash defined as one complete movement from maximum to minimum amplitude and back) following 10 min exposure to the drug. Ethanol was mixed with Dent's solution at a concentration of 400 mM, which has previously been shown to produce a ~70% reduction in locomotion rate in wild-type worms^{22,25}.

Bioinformatics

Secondary analysis

To gain a better insight into the biological pathways involved in the link between alcohol consumption and liver damage, we used the genetic variants within the genes highlighted by our model organism experiments and performed a series of secondary analyses using human data. We explored publicly available data from the UK Biobank deposited in the Edinburgh Gene Atlas²⁶ using Phewas (Phenome-wide association analysis) databases to obtain association results between the genetic variants and 778 traits. We additionally used the genes highlighted by our model organism experiments to assess the causal effect of gene expression on liver conditions. Within these genes, the SNPs that have been identified to have a *cis*-eQTL effect within the previously published studies were selected and used as MR instrument against liver conditions within the *twosampleMR* package in R.

Statistical analyses

Within the Airwave study sample, we performed a linear regression to study the association of alcohol consumption with each of the metabolomic features (Metabolome-wide association study; MWAS). We adjusted the statistical analysis for age, sex, smoking status, and salary class. To account for multiple testing and the high degree of correlation in metabolomics datasets, we used a permutation-based method to estimate the significance level of the associations^{27,28}. For each metabolomics platform, a *P*-value threshold equivalent to adjusting to a 5% Family-Wise Error Rate (i.e., Bonferroni method) was computed. A series of hypergeometric tests implemented in the R package *MetaboAnalystR*²⁹ was used for pathway enrichment analysis where an FDR threshold of 0.05 was used as a significance threshold for each of the metabolomics platforms. To obtain an estimation for the association of known alcohol SNPs with our alcohol-associated metabolites to be used in the MR analysis, we performed linear regression analysis within the Airwave sample (see “Supplementary Methods” for details of the GWAS on metabolomics). Linear analyses between SNPs and metabolites were conducted for each metabolomic feature with adjustment for age, sex, and genetic principal components within a subsample of Airwave that included participants with both genetic and metabolite data (N = 1970). In *C. elegans*, locomotion rate was presented normalized as a percentage of the mean thrashing rate of untreated worms measured each day. All worm data were expressed as mean ± SE with an N = 30 individual worms. Locomotion rate significance was assessed by one-way analysis of variance (ANOVA) with post-hoc Tukey test for multiple comparisons. Statistical analyses of the *Drosophila* experimental data were performed using GraphPad Prism (www.graphpad.com). *Drosophila* data were presented as the mean values, and the error bars indicate ± SD.

Results

To understand biological effects of alcohol consumption in human population, we investigated the circulating metabolites within the Airwave study sample using an agnostic approach which revealed the association of 152 unique metabolites with alcohol consumption (Supplementary Data). Using MR approach to examine the causality of the above associations, we identified a possible causal association (Table 2) of alcohol consumption on changing circulating level for several lipid metabolites (TAGs, Diacylglycerols, Glycerophosphocholines, Sphingolipids), and an alkaloid (piperine). The most statistically significant causal association was observed with a triacylglycerol TG 60:2 ($\beta = 1.24$; 95% CI 0.52, 1.95; *P*-value = 0.002). Pathway analysis on the 152 unique alcohol-metabolite associations showed that the linoleic acid (LNA) and alpha linolenic acid (ALA) metabolism pathway (LNA/ALA) within the Small Molecule Pathway Database (SMPDB) was enriched with alcohol-associated metabolomic features (*P*_{fdr} = 5.67×10^{-3}). These features were annotated to Tetracosapentaenoic acid (24:5n-3; $\beta = 0.01$; 95% CI 0.008, 0.012; *P*-value = 4.2×10^{-12}), Eicosapentaenoic acid ($\beta = 0.009$; 95% CI 0.007, 0.011; *P*-value = 2.4×10^{-12}), Stearidonic acid ($\beta = 0.008$; 95% CI 0.006, 0.01; *P*-value = 2.2×10^{-9}), Arachidonic acid ($\beta = 0.007$; 95% CI 0.005, 0.009; *P*-value = 9.7×10^{-8}) and Adrenic acid ($\beta = 0.01$; 95% CI 0.008, 0.012; *P*-value = 9.7×10^{-8}).

Metabolite	Name	Main class	Beta (95% CI)	P-value
SLPOS_457.3339_0.7190	CAR DC18:1	Fatty esters	-1.13 (-2.14, -0.11)	0.036
SLPOS_579.5352_8.2395	DG 34:0	DAG	0.91 (0.05,1.77)	0.047
SLPOS_606.5548_8.3326	DG 36:1	DAG	0.92 (0.07,1.77)	0.041
SLPOS_745.5603_5.7424	PC 33:2	Glycerophosphocholines	0.78 (0.07,1.5)	0.040
SLPOS_629.5424_6.9217	PE 38:4	Glycerophosphoethanolamines	1.2 (0.15,2.25)	0.032
SHPOS_625.5174_2.8914	PE 38:5	Glycerophosphoethanolamines	1.1 (0.1,2.1)	0.039
SLPOS_201.0516_0.6580	Piperine	Alkaloids	1.16 (0.1,2.22)	0.040
SHPOS_807.6348_4.4974	SM 40:2;O2	Sphingomyelins	1.06 (0.15,1.96)	0.028
SHPOS_827.7093_0.5837	TG 48:1	TAG	0.95 (0.13,1.77)	0.030
SLPOS_898.7965_11.2257	TG 53:1	TAG	0.93 (0.1,1.75)	0.036
SLPOS_888.8079_10.7574	TG 53:3	TAG	0.96 (0.06,1.85)	0.045
SLPOS_607.5599_11.1277	TG 54:2	TAG	0.95 (0.19,1.71)	0.019
SLPOS_1027.7392_10.9051	TG 54:3	TAG	1.06 (0.21,1.91)	0.020
SLPOS_984.8508_11.7812	TG 58:1	TAG	1.12 (0.44,1.81)	0.003
SLPOS_687.6307_11.5996	TG 58:2	TAG	1.08 (0.27,1.9)	0.014
SLPOS_689.6461_11.7895	TG 60:2	TAG	1.24 (0.52,1.95)	0.002

Table 2. Overview of the causal effect of alcohol on circulating metabolites using the inverse variance weighted two-sample Mendelian randomization multiple instrument method. Effect estimates and 95% CI is given for the inverse variance weighted method of Mendelian randomization. CI confidence interval, DAG diradylglycerols, TAG triradylglycerols.

We used *Drosophila hppy* mutant, described to have an increased resistance to ethanol sedation³⁰, as positive control in all fly experiments. *Drosophila* with knock-down of *ARPC1B* (*arp1*), *GPN1* (*CG3704*), *WDPCP* (*frz*), *MLXIPL* (*mondo*), *SCN8A* (*para*), and *TENM2* (*Ten-m*) together with *hppy* mutant were exposed to food supplemented with 15% Ethanol in a CAFE assay. A significant difference (Fig. 2A,B) was observed between ethanol consumed (μL) by flies with RNAi knockdown of *TENM2* (*Ten-m*). Following exposure to ethanol vapour, the effect on Sedation Time 50% (ST50; minute) was quantified and we observed that in comparison to control (*hppy*), RNAi knockdown of fly orthologues of *WDPCP* (*frz*) showed a faster rate of sedation whilst *TENM2* (*Ten-m*), *GPN1* (*CG3704*), *ARPC1B* (*arp1*) and *SCN8A* (*para*) showed a slower rate of sedation indicated by a higher ST50. *WDPCP* (*frz*), *TENM2* (*Ten-m*) and *GPN1* (*CG3704*) knockdown flies showed reduced TAG levels (Fig. 2C).

C. elegans loss-of-function mutants of *TENM2* (*ten-1*), *KIF26A* (*vab-8*), *SLC4A8* (*abts-1*) and *SCN8A* (*cca-1*) showed significant differences in basal locomotion rate (Supplementary Fig. S1A). After acute exposure to ethanol in *C. elegans* loss-of-function mutants, significant differences were identified in normalized locomotion rate of *ACTR1B* (*arp-1*), *ARPC1B* (*arx-3*) and *MAPT* (*ptl-1*) in comparison with Bristol N2 wild-type worms (Fig. 3A). RNAi knockdown of these genes confirmed that in comparison to controls, RNAi knockdown of worm orthologues of *ACTR1B* (*arp-1*) and *MAPT* (*ptl-1*) and *ARPC1B* (*arx-3*) did not have any effect on basal locomotion rate (Supplementary Fig. S1B) but RNAi knockdown of *ACTR1B* (*arp-1*) and *MAPT* (*ptl-1*) phenocopied the loss-of-function mutations after exposure to ethanol (Fig. 3B).

Secondary analysis

The Phewas analysis (Table 3) on the three genetic variants within *WDPCP*, *TENM2*, and *GPN1* showed a link between rs10078588 (*TENM2*) and food and liquid intake (beef, oily fish, fresh fruit, bread, alcohol). rs13032049 (*WDPCP*) was linked to salt intake and smoking. rs2178197 (*GPN1*) was linked to hypertension, and hematologic traits. All three SNPs showed association with disorders of lipoprotein metabolism and other lipidaemias (ICD code E78) within the UK Biobank (Table 4). We finally performed a Mendelian randomization analysis on the expression of *WDPCP* gene (ENSG00000115507) and liver conditions using cis-eQTL data (Table 5) and identified a link between expression of *WDPCP* with liver fibrosis and cirrhosis ($\beta = -0.20$; 95% CI $-0.39, -0.01$; P -value = 0.04) as well as liver and bile duct cancer ($\beta = 0.0003$; 95% CI $3.27 \times 10^{-05}, 5.85 \times 10^{-04}$; P -value = 0.02). We also observed a suggestive link with the changes in liver fatty acid-binding protein (Table 5).

Discussion

In this study we used data from humans, *C. elegans*, and *Drosophila* and identified a link between genes implicated in alcohol consumption and lipid metabolism. We identified that alcohol consumption changes the metabolites within linoleic acid (LNA) and alpha linolenic acid (ALA) metabolism pathway (LNA/ALA) and demonstrates causal effect on changes in several lipid metabolites. We highlighted that change of function of the genes implicated in alcohol consumption leads to changes in ethanol consumption, sedation after exposure to ethanol vapor, and changes in accumulation of fat in *Drosophila* as well as changes in locomotion rate after exposure to ethanol in *C. elegans*. Our results demonstrate that three alcohol-implicated genes namely *WDPCP* (*frz*), *TENM2* (*Ten-m*), and *GPN1* (*CG3704*) might be involved in fat accumulation.

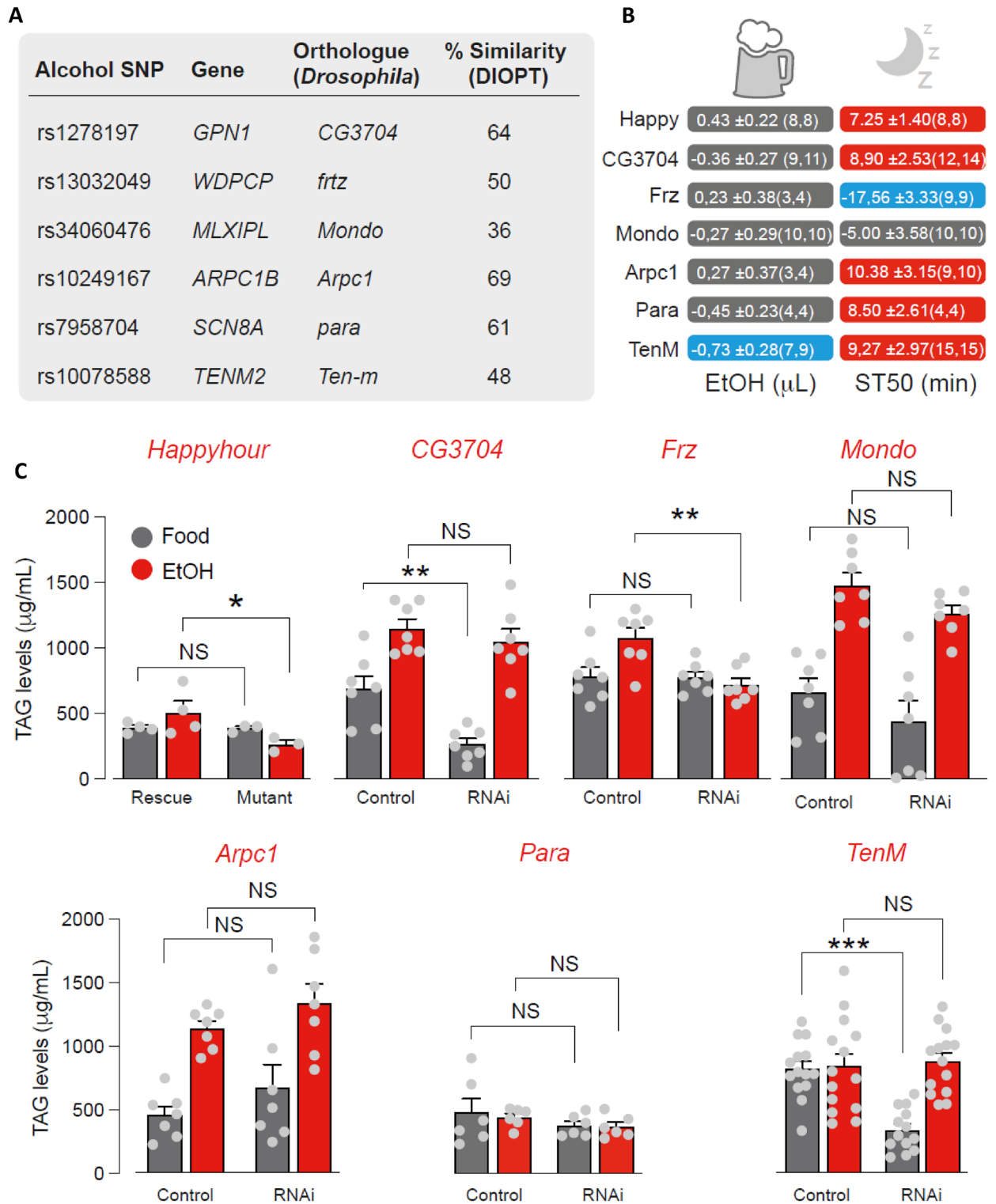


Figure 2. Analysis of alcohol intake and sedation in adult flies. (A) Mapping *Drosophila* orthologues of human genes involved in ethanol consumption. Similarity of protein alignment based on the DRSC integrative ortholog prediction tool (DIOPT). (B) Analysis of ethanol intake (left column) and sedation (right column) in the *Drosophila* RNAi lines. The numerical values show the difference between means ± standard error of mean (RNAi–Control). Red or blue correspond to the measurements that were significantly increased or decreased, respectively ($P \leq 0.05$) whereas grey indicates a significance higher than 0.05. The statistical significance was determined using an unpaired t-test. Values in parenthesis are the number of biological replicates for respectively, the control and the RNAi line. (C) Analysis of TAG levels in adult flies fed either with normal or ethanol-containing food (mean ± standard error of mean; asterisks, 2-way ANOVA with Tukey’s multiple comparisons test). The number of biological replicates per experimental variable (n) is indicated in either the respective figure or figure legend. No sample was excluded from the analysis unless otherwise stated. Blinding was not performed. Normality was assessed before deciding on which parametric or non-parametric test to use for inferential statistics. Statistical significance is indicated as * for $P < 0.05$, ** for $P < 0.01$, *** for $P < 0.001$, **** for $P < 0.0001$ and NS for $P \geq 0.05$.

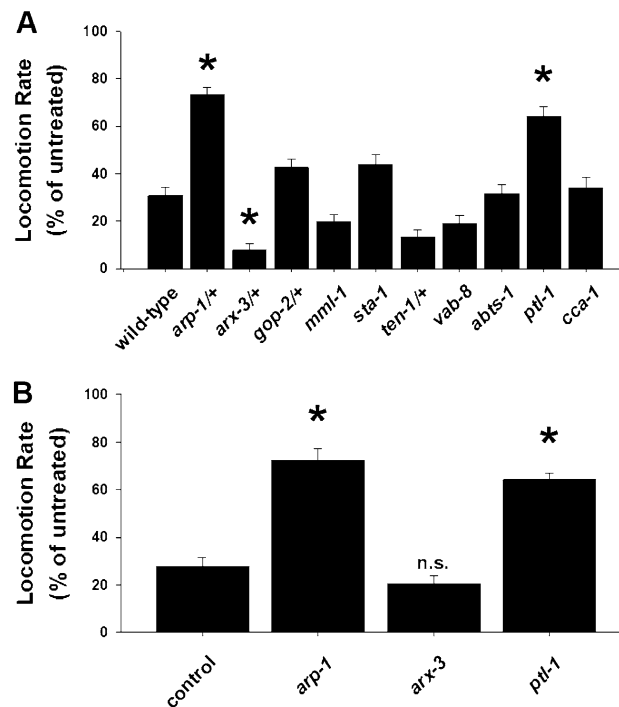


Figure 3. Quantification of alcohol phenotypes for *C. elegans* genes. (A) Nematodes with *loss-of-function* mutations in worm orthologues of *ACTR1B* (*arp-1/+*), *ARPC1B* (*arx-3/+*), *GPN1* (*gop-2/+*), *MLXIPL* (*mml-1*), *STAT6* (*sta-1*), *TENM2* (*ten-1/+*), *KIF26A* (*vab-8*), *SLC4A8* (*abts-1*), *MAPT* (*ptt-1*) and *SCN8A* (*cca-1*) were acutely exposed to ethanol and the resultant effect on rate of locomotion (thrashes per minute) was quantified. In comparison with Bristol N2 wild-type worms, significant differences were identified for *arp-1*, *arx-3* and *ptt-1*. Data is presented normalized to locomotion rate of untreated worms. * $P < 0.01$. (B) RNAi confirmation of positively identified genes involved in alcohol phenotypes. In comparison to controls, RNAi knockdown of worm orthologues of *ACTR1B* (*arp-1*) and *MAPT* (*ptt-1*) phenocopied the *loss-of-function* mutations, whereas *ARPC1B* (*arx-3*) knockdown had no effect. * $P < 0.01$. n.s., not significant.

In our study, we identified several metabolites to be causally altered due to alcohol consumption and identified the LNA/ALA pathway involved in alcohol consumption. A metabolite, eicosapentaenoic acid that is synthesized from linolenic acid, has previously been associated with altered acute behavioral responses to alcohol in *C. elegans* (altered locomotion)³¹, and mice³². In addition in humans, long-chain polyunsaturated fatty acids are shown to be associated with alcohol sensitivity³³ and genes that are essential to generate ω -3 long-chain polyunsaturated fatty acids are shown to be associated with alcohol-related phenotypes³⁴. Previous studies also showed associations between lipid metabolites and alcohol consumption³⁵. We previously showed alteration of high-density lipoprotein to be associated with alcohol consumption within the UK Biobank¹⁰. The metabolomics part of the current study is a step forward in that firstly it uses causal inference tools (MR method) to demonstrate the causality of the association between lipid metabolites specifically long chain fatty acids with alcohol consumption. Secondly, by performing a pathway analysis, we were able to pinpoint a specific LNA/ALA pathway in alcohol consumption.

In our study, we used alcohol-associated genetic variants to explore (1) the alcohol-induced biological changes in human metabolites and (2) alcohol-induced biological effect of genes annotated to alcohol-associated genetic variants in *C. elegans* and *Drosophila melanogaster*. Of the alcohol-implicated genes that we investigated in *C. elegans*, *ACTR1B* (*arp-1*) and *MAPT* (*ptt-1*), show significant effects on the worms' locomotion upon acute exposure to ethanol. In addition, *TENM2* (*Ten-m*) shows significant effects on ethanol consumption in *Drosophila* and apart from *MLXIPL* (*mondo*), RNAi lines for all the genes investigated in *Drosophila* change the time to sedation from ethanol.

WDPCP which was first characterized in *Drosophila*³⁶ is known to be involved in cell polarity and ciliogenesis^{37,38}. In humans, mutations in *WDPCP* gene cause Bardet-Biedl Syndrome³⁹ presenting with a variety of symptoms in different organs including obesity, blindness, and polydactyly. *WDPCP* gene is widely expressed known to be involved in hedgehog signaling⁴⁰. Recently, Liu and colleagues¹³ identified that a G allele in rs13032049 within the *WDPCP* gene was associated with an increased consumption of alcohol. The G allele in rs13032049 shows a strong association with a lower expression of the *WDPCP* gene. We found the SNP to be strongly associated with increased liver enzyme GGT and behavioral traits such as smoking and adding salt to food. In our *Drosophila* experiments, the observed effect of *WDPCP* (*frz*) knockdown on the changed TAG levels occurred under exposure to ethanol and followed similar patterns as *hppy* mutants. This implies that minimizing alcohol consumption could reduce fat accumulation and thus could potentially reduce the risk of hepatic lipogenesis. In our further MR analysis, we used publicly available databases derived from GWAS in human population and showed a link between the gene expression of *WDPCP* and liver fibrosis and liver cirrhosis. The

Trait	Beta	Z value	P-value
rs10078588-A (<i>TENM2</i>)			
Bread intake, number of slices/week	0.09	5.66	7.57×10^{-9}
Alcohol weekly intake frequency (high to low frequency)	-0.02	5.21	9.63×10^{-8}
Oily fish intake, number of times/week	-0.01	4.53	2.88×10^{-6}
Beef intake, number of intake times/week	0.01	4.45	4.34×10^{-6}
Fresh fruit intake, number of pieces/day	-0.01	4.44	4.45×10^{-6}
rs13032049-G (<i>WDPCP</i>)			
Salt added to food, frequency	0.01	5.01	2.77×10^{-7}
Smoking status, current vs. previous and never	0.01	4.12	1.88×10^{-5}
rs2178197-G (<i>GPN1</i>)			
Hypertension	-3.60×10^{-3}	-4.26	1.02×10^{-5}
Mean spheroid cell volume, fl	0.04	-4.98	3.15×10^{-7}
Red blood cell (erythrocyte) count, $\times 10^{12}$ cells/L	-3.30×10^{-3}	-5.41	3.21×10^{-8}
Number of treatments/medications taken	-0.02	-4.64	1.71×10^{-6}
Monocyte percentage, %	0.02	-4.97	3.39×10^{-7}
Number of operations, self-reported	-0.01	-4.33	7.58×10^{-6}
Mean reticulocyte volume, fL	0.06	-4.39	5.66×10^{-6}
High light scatter reticulocyte percentage, %	-1.60×10^{-3}	-4.37	6.32×10^{-6}
Reticulocyte count, $\times 10^{12}$ cells/L	-2.00×10^{-4}	-5.15	1.33×10^{-7}
Reticulocyte percentage, %	-4.20×10^{-3}	-4.09	2.12×10^{-5}
Lymphocyte percentage, %	0.06	-4.49	3.59×10^{-6}
Neutrophil percentage, %	-0.07	-4.48	3.81×10^{-6}
High light scatter reticulocyte count, $\times 10^{12}$ cells/L	-1.00×10^{-4}	-5.39	3.50×10^{-8}

Table 3. Overview of the significant associations between SNPs in *WDPCP*, *TENM2*, and *GPN1* with Phewas traits within the UK Biobank Edinburgh Gene Atlas. Significant associations are depicted in bold. Statistical significance level P -value = 2.14×10^{-5} was considered equivalent to a P -value adjusted for multiple testing for the analysis of 3 SNPs and 778 phenotypes within the Gene Atlas. The associations are provided per copy of the allele that increases alcohol intake.

SNP	chromosome	Base pair location	Effect allele	Beta (effect estimate)	95% CI lower bound	95% CI upper bound	P-value	MAF	HWE
rs10078588	5 (<i>TENM2</i>)	166,816,176	A	-0.001	-2.49×10^{-04}	-2.44×10^{-03}	0.02	0.47	0.5
rs13032049	2 (<i>WDPCP</i>)	63,581,507	G	-0.001	-1.59^{-04}	-2.60×10^{-03}	0.03	0.28	0.51
rs2178197	2 (<i>GPN1</i>)	27,860,551	G	-0.001	-3.36×10^{-04}	-2.55×10^{-03}	0.01	0.42	0.21

Table 4. Overview of the association of SNPs in *WDPCP*, *TENM2*, and *GPN1* with disorders of lipoprotein metabolism and other lipidaemias (ICD code E78) within the UK Biobank Edinburgh Gene Atlas. Results obtained from the Edinburgh Gene Atlas. *CI* confidence interval, *SNP* single nucleotide polymorphism, *MAF* minor allele frequency, *HWE* Hardy Weinberg Equilibrium.

Liver trait	Number of SNPs	Effect estimate	Standard error	P value
Fatty acid-binding protein, liver	6	-0.10	0.05	0.06
Liver enzyme levels (alanine transaminase)	6	-0.002	0.001	0.2
Fibrosis and cirrhosis of liver	6	-0.20	0.1	0.04
Liver & bile duct cancer	6	0.0003	0.0001	0.03

Table 5. Overview of the significant results from inverse variance weighted Mendelian randomization analysis for the effect of gene expression of ENSG00000115507 on liver traits using the MRC IEU OpenGWAS data infrastructure⁵¹.

analysis also showed a suggestive link with the liver fatty acid binding protein that is involved in the metabolism of lipids⁴¹. This evidence suggests that *WDPCP* might be an important gene involved in the pathway between alcohol consumption, accumulation of fat and liver fibrosis. This could have public health implications in terms of the identification of high-risk groups and targeting preventive measures as well as drug development. More

studies in vivo and in vitro are needed to focus on *WDPCP* and provide more details on its role in lipid metabolism and liver pathologies.

Changes in TAG levels of *Drosophila* occurred with exposure to normal food rather than ethanol for RNAi knockdown of *TENM2* (*Ten-m*) and *GPN1* (*CG3704*) indicating that loss of function of these genes could have a direct role in the accumulation of fat in liver independent of exposure to ethanol. RNAi knockdown of both genes shows increased tolerance to the sedative effect of ethanol which could justify the effect of these genes on a more frequent alcohol consumption in humans, possibly due to alcohol tolerance. *GPN1* is located on chromosome 2p23.3 and the encoded protein is implicated in the regulation of TGF β superfamily signaling⁴² that is demonstrated to play a role in obesity⁴³, accumulation of fat in the liver^{42,44} as well as regulation of *ADH1* gene that enhances alcohol-induced liver damage and lipid metabolism⁴⁵. The existing evidence alongside our findings on the role of *GPN1* in alcohol consumption and lipid metabolism in *Drosophila* implies that *GPN1* might play a role upstream of TGF β in the regulation of the metabolism of alcohol and lipids. Further studies are needed to highlight the relationship between *GPN1* and TGF β in alcohol consumption and alcohol-induced liver damage.

TENM2 (*Ten-m*) is located on chromosome 5q34, and the encoded protein is involved in cell adhesion⁴⁶. *TENM2* is found to be highly enriched in white adipocyte progenitor cells⁴⁷. *TENM2* deficiency in human fat cells leads to expression of UCPI, the primary marker of brown adipose tissue⁴⁸. Genetic variants in *TENM2* have shown to be linked to obesity⁴⁹. Our secondary analyses confirmed an association between the genetic variant in *TENM2* and excess of food and liquid intake which suggests the link between *TENM2*, and alcohol consumption could also be due to systematic increase in consumption of all food and beverages rather than alcohol alone. The evidence in this study alongside the existing literature highlights that the observed effect of *TENM2* (*Ten-m*) RNAi knockdown on the changes in TAG levels in *Drosophila* could potentially be related to biological pathways implicated in adipose tissue rather than pure liver-related pathways.

One strength of our study is in that we performed our analyses in human and two different model organisms, allowing for a more comprehensive insight into biological mechanisms involved in the function of alcohol consumption genes under different biological scenarios. A second strength of this study is in the use of RNAi technique which provides insight into the function of genes and what biological manifestation they would have when exposed to ethanol. The third strength of our study is in the use of CAFE assay that allows for the investigation of food and alcohol consumption in *Drosophila* in a more controlled environment. In our CAFE assay, each experimental box per genotype contained both normal food and ethanol food (food supplemented with 15% ethanol) providing the insects with a choice. Another strength of our study is that in our TAG levels experiments, we made our conclusions based on the comparisons between flies (RNAi vs. control) that were exposed to identical food and environmental conditions which increases the robustness of our conclusions. Finally, we combined the results with human studies to get better insight into the link between alcohol consumption and lipid metabolites.

Alcohol consumption in Airwave participants was calculated based on self-reported data which could have affected the precision of the alcohol consumed due to recall bias. To reduce this limitation, the duration of recall was limited to the last 7 days in the Airwave study. We should acknowledge that the concentration of alcohol within alcoholic drinks is not standard⁵⁰, and our calculation of alcohol consumed could be affected by these variations. Although the genetic variants used for our investigations were originally found in human studies and we also performed a metabolomics analysis between alcohol consumption and circulating metabolites in the human population, the main part of the study was performed in model organisms and the results of this study might not directly generalisable to patients and the public without performing further population studies.

Conclusion

We found that alcohol-associated genes may be involved in the metabolism of lipids. Our study highlights three genes (*WDPCP*, *TENM2*, and *GPN1*) that may be involved in the accumulation of lipids. Of these genes, *WDPCP* exhibits its effects on lipid accumulation in *Drosophila* with exposure to ethanol. The gene expression of *WDPCP* in the human population supports a link to liver fibrosis. Further studies are necessary to investigate the role of this gene in ALD.

Data availability

All data generated or analyzed during this study are included in this published article (and its Supplementary Information files).

Received: 16 June 2023; Accepted: 13 November 2023

Published online: 23 November 2023

References

1. World Health, O. *Global Status Report on Alcohol and Health 2018: Executive Summary* (World Health Organization, 2018).
2. Osna, N. A., Donohue, T. M. Jr. & Kharbanda, K. K. Alcoholic liver disease: Pathogenesis and current management. *Alcohol Res.* **38**, 147–161 (2017).
3. Evangelou, E. *et al.* New alcohol-related genes suggest shared genetic mechanisms with neuropsychiatric disorders. *Nat. Hum. Behav.* **3**, 950–961 (2019).
4. Pazoki, R. *et al.* Genetic analysis in European ancestry individuals identifies 517 loci associated with liver enzymes. *Nat. Commun.* **12**, 2579 (2021).
5. Elliott, P. *et al.* The Airwave Health Monitoring Study of police officers and staff in Great Britain: Rationale, design and methods. *Environ. Res.* **134**, 280–285 (2014).
6. Kanehisa, M. & Goto, S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **28**, 27–30 (2000).
7. Jewison, T. *et al.* SMPDB 2.0: Big improvements to the small molecule pathway database. *Nucleic Acids Res.* **42**, D478–D484 (2014).
8. Hastie T, T. R., Narasimhan, B. & Chu, G. *impute: Imputation for Microarray Data* (2020).

9. Roshchupkin, G. V. *et al.* HASE: Framework for efficient high-dimensional association analyses. *Sci. Rep.* **6**, 36076 (2016).
10. O'Farrell, E., Jiang, X., Aljifri, S. & Pazoki, R. Molecular alterations caused by alcohol consumption in the UK biobank: A Mendelian randomisation study. *Nutrients* **14**, 2943 (2022).
11. Smith, G. D. & Ebrahim, S. "Mendelian randomization": Can genetic epidemiology contribute to understanding environmental determinants of disease?. *Int. J. Epidemiol.* **32**, 1–22 (2003).
12. Jiang, X., Anasanti, M. D., Drenos, F., Blakemore, A. I. & Pazoki, R. Urinary sodium excretion enhances the effect of alcohol on blood pressure. *Healthcare* **10**, 1296 (2022).
13. Liu, M. *et al.* Association studies of up to 12 million individuals yield new insights into the genetic etiology of tobacco and alcohol use. *Nat. Genet.* **51**, 237–244 (2019).
14. Hoggart, C. J., Clark, T. G., De Iorio, M., Whittaker, J. C. & Balding, D. J. Genome-wide significance for dense SNP and resequencing data. *Genet. Epidemiol.* **32**, 179–185 (2008).
15. The Genotype-Tissue Expression (GTEx) project. *Nat Genet* **45**, 580–585 (2013).
16. Thompson, A., *et al.* Functional validity, role, and implications of heavy alcohol consumption genetic loci. *Science advances* **6**, eaay5034 (2020).
17. Pandey, P., Singh, A., Kaur, H., Ghosh-Roy, A. & Babu, K. Increased dopaminergic neurotransmission results in ethanol dependent sedative behaviors in *Caenorhabditis elegans*. *PLoS Genet.* **17**, e1009346 (2021).
18. Alaimo, J. T. *et al.* Ethanol metabolism and osmolarity modify behavioral responses to ethanol in *C. elegans*. *Alcoholism* **36**, 1840–1850 (2012).
19. Diegelmann, S. *et al.* The CApillary FEeder assay measures food intake in *Drosophila melanogaster*. *J. Vis. Exp.* **121**, 15024 (2017).
20. Aleyakpo, B. *et al.* G-protein aq gene expression plays a role in alcohol tolerance in *Drosophila melanogaster*. *Brain Neurosci. Adv.* **3**, 2398212819883081 (2019).
21. Tennessen, J. M., Barry, W. E., Cox, J. & Thummel, C. S. Methods for studying metabolism in *Drosophila*. *Methods* **68**, 105–115 (2014).
22. Johnson, J. R., Rajamanoharan, D., McCue, H. V., Rankin, K. & Barclay, J. W. Small heat shock proteins are novel common determinants of alcohol and nicotine sensitivity in *Caenorhabditis elegans*. *Genetics* **202**, 1013–1027 (2016).
23. Kamath, R. S. & Ahringer, J. Genome-wide RNAi screening in *Caenorhabditis elegans*. *Methods* **30**, 313–321 (2003).
24. Rual, J. F. *et al.* Toward improving *Caenorhabditis elegans* phenome mapping with an ORFeome-based RNAi library. *Genome Res.* **14**, 2162–2168 (2004).
25. Davies, A. G. *et al.* A central role of the BK potassium channel in behavioral responses to ethanol in *C. elegans*. *Cell* **115**, 655–666 (2003).
26. Canela-Xandri, O., Rawlik, K. & Tenesa, A. An atlas of genetic associations in UK Biobank. *Nat. Genet.* **50**, 1593–1599 (2018).
27. Castagné, R. *et al.* Improving visualization and interpretation of metabolome-wide association studies: An application in a population-based cohort using untargeted 1H NMR metabolic profiling. *J. Proteome Res.* **16**, 3623–3633 (2017).
28. Chadeau-Hyam, M. *et al.* Metabolic profiling and the metabolome-wide association study: Significance level for biomarker identification. *J. Proteome Res.* **9**, 4620–4627 (2010).
29. Chong, J. *et al.* MetaboAnalyst 40: Towards more transparent and integrative metabolomics analysis. *Nucleic Acids Res.* **46**, W486–W494 (2018).
30. Corl, A. B. *et al.* Happyhour, a Ste20 family kinase, implicates EGFR signaling in ethanol-induced behaviors. *Cell* **137**, 949–960 (2009).
31. Raabe, R. C., Mathies, L. D., Davies, A. G. & Bettinger, J. C. The omega-3 fatty acid eicosapentaenoic acid is required for normal alcohol response behaviors in *C. elegans*. *PLoS ONE* **9**, e105999 (2014).
32. Wolstenholme, J. T. *et al.* Dietary omega-3 fatty acids differentially impact acute ethanol-responsive behaviors and ethanol consumption in DBA/2J versus C57BL/6J mice. *Alcoholism* **42**, 1476–1485 (2018).
33. Edwards, A. C. *et al.* Long-chain ω -3 levels are associated with increased alcohol sensitivity in a population-based sample of adolescents. *Alcoholism* **43**, 2620–2626 (2019).
34. Aliev, F., Barr, P. B., Davies, A. G., Dick, D. M. & Bettinger, J. C. Genes regulating levels of ω -3 long-chain polyunsaturated fatty acids are associated with alcohol use disorder and consumption, and broader externalizing behavior in humans. *Alcoholism* **46**, 1657–1664 (2022).
35. Langenau, J., Boeing, H., Bergmann, M. M., Nöthlings, U. & Oluwabemigun, K. The association between alcohol consumption and serum metabolites and the modifying effect of smoking. *Nutrients* **11**, 2331 (2019).
36. Collier, S., Lee, H., Burgess, R. & Adler, P. The WD40 repeat protein fritz links cytoskeletal planar polarity to frizzled subcellular localization in the drosophila epidermis. *Genetics* **169**, 2035–2045 (2005).
37. Cui, C. *et al.* Wdpcp, a PCP protein required for ciliogenesis, regulates directional cell migration and cell polarity by direct modulation of the actin cytoskeleton. *PLoS Biol.* **11**, e1001720 (2013).
38. Kim, S. K. *et al.* Planar cell polarity acts through septins to control collective cell movement and ciliogenesis. *Science* **329**, 1337–1340 (2010).
39. Forsythe, E. & Beales, P. L. Bardet-Biedl syndrome. *Eur. J. Hum. Genet.* **21**, 8–13 (2013).
40. Langhans, M. T. *et al.* Wdpcp regulates cellular proliferation and differentiation in the developing limb via hedgehog signaling. *BMC Dev. Biol.* **21**, 10 (2021).
41. Martin, G. G. *et al.* Decreased liver fatty acid binding capacity and altered liver lipid distribution in mice lacking the liver fatty acid-binding protein gene. *J. Biol. Chem.* **278**, 21429–21438 (2003).
42. Taneja-Bageshwar, S. & Gumienny, T. L. Regulation of TGF β superfamily signaling by two separable domains of glypican LON-2 in *C. elegans*. *Worm* **2**, e23843 (2013).
43. Zhao, J. *et al.* Hepatocyte TGF- β signaling inhibiting WAT browning to promote NAFLD and obesity is associated with Let-7b-5p. *Hepatol. Commun.* **6**, 1301–1321 (2022).
44. Xiang, X. *et al.* Impaired reciprocal regulation between SIRT6 and TGF- β signaling in fatty liver. *FASEB J.* **36**, e22335 (2022).
45. Ciucan, L. *et al.* TGF-beta enhances alcohol dependent hepatocyte damage via down-regulation of alcohol dehydrogenase I. *J. Hepatol.* **52**, 407–416 (2010).
46. Boucard, A. A., Maxeiner, S. & Südhof, T. C. Latrophilins function as heterophilic cell-adhesion molecules by binding to teneurins: Regulation by alternative splicing. *J. Biol. Chem.* **289**, 387–402 (2014).
47. Tews, D. *et al.* Comparative gene array analysis of progenitor cells from human paired deep neck and subcutaneous adipose tissue. *Mol. Cell. Endocrinol.* **395**, 41–50 (2014).
48. Tews, D. *et al.* Teneurin-2 (TENM2) deficiency induces UCP1 expression in differentiating human fat cells. *Mol. Cell. Endocrinol.* **443**, 106–113 (2017).
49. Scannell Bryan, M. *et al.* Genome-wide association studies and heritability estimates of body mass index related phenotypes in bangladeshi adults. *PLoS ONE* **9**, e105062 (2014).
50. Ferner, R. E. & Chambers, J. Alcohol intake: Measure for measure. *BMJ* **323**, 1439–1440 (2001).
51. Elsworth, B. *et al.* The MRC IEU OpenGWAS data infrastructure. *Biorxiv* **35**, 99 (2020).

Acknowledgements

R.P. was supported by Rutherford Fund fellowship from the Medical Research Council (MR/R026505/1 and MR/R026505/2). B.A., X.J., and F.O. were supported by Rutherford Fund from Medical Research Council MR/R026505/2. R.M. was funded by the President's PhD Scholarship from Imperial College London. PE is Director of the MRC Centre for Environment and Health and acknowledges support from the Medical Research Council (MR/S019669/1). PE also acknowledges support from the UK Dementia Research Institute, Imperial College London (UKDRI-5001), Health Data Research UK London (HDRUK-1004231) and the British Heart Foundation Imperial College London Centre for Research Excellence (BHF-RE/18/4/34215). The Airwave Health Monitoring Study was funded by the UK Home Office (780- TETRA, 2003-2018) and is currently funded by the MRC and ESRC (MR/R023484/1) with additional support from the NIHR Imperial College Biomedical Research Centre in collaboration with Imperial College NHS Healthcare Trust. R.C.P. is supported by the UK Dementia Research Institute (UKDRI-5001), which receives its funding from UK DRI Ltd, funded by the UK Medical Research Council, Alzheimer's Society and Alzheimer's Research UK. Work in LMM's laboratory is supported by the UK Medical Research Council, intramural project MC_UU_00025/3 (RG94521). The views expressed are those of the authors and not necessarily those of the sponsors. We thank Prof. Ulrike Heberlein, (Janelia Research Campus, Virginia, USA) for generously providing us the hppy17-51 fly lines. This research was funded, in whole or in part, by the Medical Research Council (MR/R026505/1 and MR/R026505/2). A CC BY or equivalent licence is applied to the Author Accepted Manuscript (AAM) arising from this submission, in accordance with the grant's open access conditions.

Author contributions

R.P. contributed to study design, supervision, leading the project, and data interpretation. X.J. performed the initial analysis including SNP selection and shortlisting for the experiments. R.C.P. performed the metabolome wide association analysis within the Airwave study, R.M. performed metabolomics GWAS within the Airwave study. F.O. performed the pathway enrichment analysis. B.A. performed the *Drosophila* experiments. S.H.Y.L. and L.M.M. supervised the *Drosophila* experiments. J.B. conducted the *C. elegans* experiments. A.B. provided scientific input and supervision regarding data metabolomics and GWAS analysis within the Airwave study. I.T. and P.E. acquired the Airwave data. F.O. and R.P. drafted the manuscript.

Competing interests

The authors declare no competing interests.


Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-023-47371-7>.

Correspondence and requests for materials should be addressed to R.P.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023