PLOS ONE

RESEARCH ARTICLE

Genetics of vegetarianism: A genome-wide association study

Nabeel R. Yaseen[1]*, Catriona L. K. Barnes[2], Lingwei Sun[3], Akiko Takeda[4], John P. Rice[3]

1 Department of Pathology, Feinberg School of Medicine, Northwestern University, Chicago, IL, United States of America, **2** Fios Genomics, Edinburgh, United Kingdom, **3** Department of Psychiatry, School of Medicine, Washington University in St. Louis, St. Louis, MO, United States of America, **4** Retired, St. Louis, MO, United States of America

* nabeel.yaseen@northwestern.edu

Check for updates

OPEN ACCESS

Citation: Yaseen NR, Barnes CLK, Sun L, Takeda A, Rice JP (2023) Genetics of vegetarianism: A genome-wide association study. PLoS ONE 18(10): e0291305. https://doi.org/10.1371/journal.pone.0291305

Editor: Qiaoping Wang, Sun Yat-Sen University, CHINA

Received: November 11, 2022

Accepted: August 19, 2023

Published: October 4, 2023

Copyright: © 2023 Yaseen et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the manuscript and its Supporting Information files.

Funding: The authors received no specific funding for this work.

Competing interests: The authors have declared that no competing interests exist.

Abstract

A substantial body of evidence points to the heritability of dietary preferences. While vegetarianism has been practiced for millennia in various societies, its practitioners remain a small minority of people worldwide, and the role of genetics in choosing a vegetarian diet is not well understood. Dietary choices involve an interplay between the physiologic effects of dietary items, their metabolism, and taste perception, all of which are strongly influenced by genetics. In this study, we used a genome-wide association study (GWAS) to identify loci associated with strict vegetarianism in UK Biobank participants. Comparing 5,324 strict vegetarians to 329,455 controls, we identified one SNP on chromosome 18 that is associated with vegetarianism at the genome-wide significant level (rs72884519, β = -0.11, $P = 4.997 \times 10^{-8}$), and an additional 201 suggestively significant variants. Four genes are associated with rs72884519: *TMEM241*, *RIOK3*, *NPC1*, and *RMC1*. Using the Functional Mapping and Annotation (FUMA) platform and the Multi-marker Analysis of GenoMic Annotation (MAGMA) tool, we identified 34 genes with a possible role in vegetarianism, 3 of which are GWAS-significant based on gene-level analysis: *RIOK3*, *RMC1*, and *NPC1*. Several of the genes associated with vegetarianism, including *TMEM241*, *NPC1*, and *RMC1*, have important functions in lipid metabolism and brain function, raising the possibility that differences in lipid metabolism and their effects on the brain may underlie the ability to subsist on a vegetarian diet. These results support a role for genetics in choosing a vegetarian diet and open the door to future studies aimed at further elucidating the physiologic pathways involved in vegetarianism.

Introduction

Abstention from the consumption of animal flesh has been advocated for thousands of years for religious, ethical, environmental, and health-related reasons [1–8]. Eastern religious traditions that discourage meat consumption include Hinduism and Buddhism. In ancient Greece, vegetarianism was practiced as early as the 6th century BC by the followers of Pythagoras and Orphism [2, 4]. During the Renaissance and Enlightenment periods several prominent

personalities in Europe practiced vegetarianism, and vegetarian societies began to be established in Europe and America in the 19th century [4].

While religious and ethical considerations were, and to a large extent still are, major motivations behind adopting a vegetarian diet, recent research has provided evidence for its health benefits, including lower risk of metabolic syndrome, obesity, dyslipidemias, diabetes, cardiovascular disease, and some cancers [5, 6, 8–15]. On the other hand, there is evidence that vegetarian diets can lead to nutritional deficiencies and may be associated with negative effects such as anemia, dental erosion, osteopenia, and psychological disorders [16–29].

Although vegetarianism is increasing in popularity, vegetarians remain a small minority of people worldwide; for example, in the United States, vegetarians comprise approximately 3–4% of the population [2, 4, 6, 30]. In the United Kingdom, 2.3% of adults and 1.9% of children are vegetarian according to the National Diet and Nutrition Survey reported in 2014. The true number of strict vegetarians is likely much lower, as a large proportion (approximately 48–64%) of self-identified vegetarians report consuming fish, poultry, and/or red meat [31–33]. This suggests that the desire to adhere to a vegetarian diet is overridden by environmental and/or biological constraints and raises the question of whether all humans are capable of surviving and thriving on a long-term strict vegetarian diet.

In support of a biological underpinning for vegetarianism, a large body of evidence points to a genetic component for food choice, including preferences for meat or vegetables as well as "healthy" or "unhealthy" eating patterns [34–50]. These findings indicate that differences in dietary preferences and/or needs among individuals are, at least to some extent, dictated by genetics.

To explore the contribution of genetics to vegetarianism, we utilized dietary and genetic data from the UK Biobank (https://www.ukbiobank.ac.uk/). The UK Biobank project is a prospective cohort study of approximately 500,000 individuals from the United Kingdom, aged between 40 and 69 at recruitment, that provides extensive genetic and phenotypic data, including comprehensive dietary information. Two categories of detailed dietary data were obtained from participants in the UK Biobank: (i) an initial touchscreen dietary questionnaire was administered at the recruitment session and 3 subsequent visits asking participants about their average intake of various food items over the previous year, and (ii) an online questionnaire based on 24-hour dietary recall of the previous day was administered towards the end of the recruitment session and at 4 additional instances thereafter, with invitations being emailed to participants at 3–4 monthly intervals. Genetic data were obtained from participants using two very similar genotyping arrays: a subset of about 50,000 participants were genotyped using the UK BiLEVE Axiom Array, while the vast majority of participants were genotyped using the closely related UK Biobank Axiom Array [51]. In this study we used genome-wide association (GWAS) to identify genetic loci that are associated with long-term adherence to a strict vegetarian diet among UK Biobank participants.

Methods

Participants

UK Biobank (UKB) is a population-based health research resource consisting of approximately 500,000 people, aged between 40 and 69 at recruitment. Phenotypic information on each participant was gathered from physical and cognitive measurements, sample collection (blood, urine, saliva), and questionnaires querying socio-demographic, lifestyle, and health-related factors. Detailed information from two questionnaires (a touchscreen questionnaire and diet by 24-hour recall questionnaire) was obtained from UKB to screen participants for inclusion in this GWAS as a vegetarianism case or control.

Ethics statement

UKB has approval from the North West Multi-centre Research Ethics Committee (MREC) as a Research Tissue Bank (RTB). This approval means that researchers do not require separate ethical clearance and can operate under the RTB approval. The data was fully anonymized by UKB prior to our accessing them. At enrollment, electronic signed consent was obtained by UKB from all participants to use their anonymized data and samples for any health-related research and they were advised that they have the right to withdraw at any time without giving a reason and without penalty [52].

Genotyping, imputation, and quality control

Genotyping, initial quality control (QC), phasing and imputation were conducted by UKB. The UKB QC pipeline was specifically designed to accommodate a large-scale data set containing ethnically diverse participants who were genotyped in many batches across two slightly different novel genotyping arrays, UK BiLEVE and UKB Axiom, which have over 95% common content. The genotyping, quality control and imputation of this data has been extensively described previously [51].

UK Biobank (UKB) carries out extensive sample and SNP quality control prior to imputation, which includes testing for batch effects, plate effects, extreme Hardy-Weinberg equilibrium departures, sex effects, array effects and discordance across control replicates. More information on the quality control carried out by UKB can be found in the Genotyping and quality control of UK Biobank documentation.

Pre-phasing was carried out by UK Biobank in order to derive the underlying haplotypes for each individual followed by imputation to estimate the unobserved genotypes [51, 53]. Imputation was carried out using the Haplotype Reference Consortium (HRC) data [54] as well as a merged panel consisting of the UK10K haplotype reference panel merged together with the 1000 Genomes Phase 3 reference panel [55]. Imputation was carried out with the IMPUTE4 program, which is a re-coded version of the haploid imputation functionality implemented in IMPUTE2 [51, 53]. The result of the imputation process is a dataset with 93,095,623 autosomal SNPs, short indels and large structural variants in 487,442 individuals and an additional 3,963,705 markers on the X chromosome.

In addition to the pre-imputation quality control checks, UKB also calculates metrics and provides data for a number of key variables, including relatedness, ethnicity, genetic and reported sex, heterozygosity, missingness, and genetic variables such as minor allele frequency (MAF) and INFO scores. UKB does not exclude samples based on these metrics, but provides either the metrics or a list of recommended samples to exclude based on these measurements; it is then up to the researcher to exclude any further samples. After receiving the data from UKB, additional in-house sample QC and SNP QC were then performed based on these key sample quality control metrics. The following is a summary of the Sample QC and SNP QC performed in-house; a more detailed description with figures is presented in Supporting Information files S1 and S2 Appendices, respectively.

Sample QC. Five sample QC metrics were assessed to produce a list of samples to exclude before downstream data analysis was performed. UK BiLEVE: samples that failed the QC procedures conducted by UKB were excluded. Samples with ethnicity other than white caucasian were excluded. To obtain an unrelated data set for analysis, individuals were excluded based on their kinship coefficient values. Any individual where the reported sex and genetic sex did not match or with sex chromosome aneuploidy was excluded. Heterozygosity and missingness outliers were excluded. Following sample QC, 161,655 unique samples were excluded from the

downstream analysis. This left a total of 340,754 samples for case/control assessment and inclusion in the GWAS.

SNP QC. Imputed genetic data were assessed using two QC metrics. Minor allele frequency (MAF): variants with a MAF below 0.01 were excluded. Variants with an INFO score below 0.7 were excluded. Following SNP QC, 83,355,424 variants were excluded from the downstream analysis. This left a total of 9,740,199 variants for inclusion in the GWAS.

Phenotype processing. The quality-controlled data set was split into vegetarianism cases and controls. Any sample that did not pass either the vegetarianism case or control criteria was not included in the GWAS. Following the vegetarianism screening process, 5,324 individuals were classed as vegetarianism cases and 329,455 individuals were classed as controls. Below is a summary of the phenotype processing; a detailed description of the phenotype processing methods is presented in S3 Appendix.

Screening vegetarianism cases. Dietary phenotype data were collected using two questionnaires: the touchscreen questionnaire, which had ~500K respondents, and the diet by 24-hour recall questionnaire, which captured ~110K of the respondents. The two questionnaires were not mutually exclusive, so both questionnaires were screened separately.

Initial exclusions were carried out based on the touchscreen questionnaire data fields asking about intake of fish, processed meat, poultry intake, beef intake, lamb/mutton intake, and pork intake. An individual was excluded if they responded "Do not know" or "Prefer not to answer" at any time point. In addition, individuals who reported eating meat within the last year and those who responded with "Prefer not to answer" were also excluded. The remaining individuals were then screened separately for the two questionnaires.

For touchscreen questionnaire screening, the data fields that were assessed in the initial exclusions were further assessed. Individuals were retained as a case if they had at least one instance within a data field answered as "Never" and every other instance was either "Never" or NA (where the individual had not completed that particular instance) within that data field. Individuals who had responded to any instance with an answer other than "Never" or NA, or had not responded to any instance within a data field were excluded.

For diet by 24-hour recall questionnaire screening, individuals who reported following a "Vegetarian" or "Vegan" diet were included. This pool of cases was then screened to exclude individuals who reported any intake of sausage, beef, pork, lamb, poultry, bacon, ham, liver, other meat, fish, seafood, and lard.

Screening the two separate questionnaires resulted in two overlapping pools of individuals classed as vegetarianism cases. Individuals who failed one questionnaire screening but passed the other were excluded from the total pool of cases. Individuals who passed the touchscreen questionnaire but did not fill out the 24-hour recall questionnaire were retained for inclusion within the analysis. This resulted in a final pool of 5,324 cases (Fig 1).

Screening vegetarianism controls. Initial vegetarianism control screening was similar to the case screening, where individuals who answered "Do not know" or "Prefer not to answer" at any time point on data fields asking about intake of fish, processed meat, poultry intake, beef intake, lamb/mutton intake, and pork intake were excluded. Remaining individuals were then screened so any individual that reported any intake on at least one occasion on any of the above data fields was retained for inclusion as a control. This resulted in a final pool of 329,455 controls.

Statistical analysis

A detailed description of the GWAS analysis is presented in S4 Appendix. Briefly, genome-wide association analysis was performed for the vegetarianism phenotype (5,324 cases and 329,455

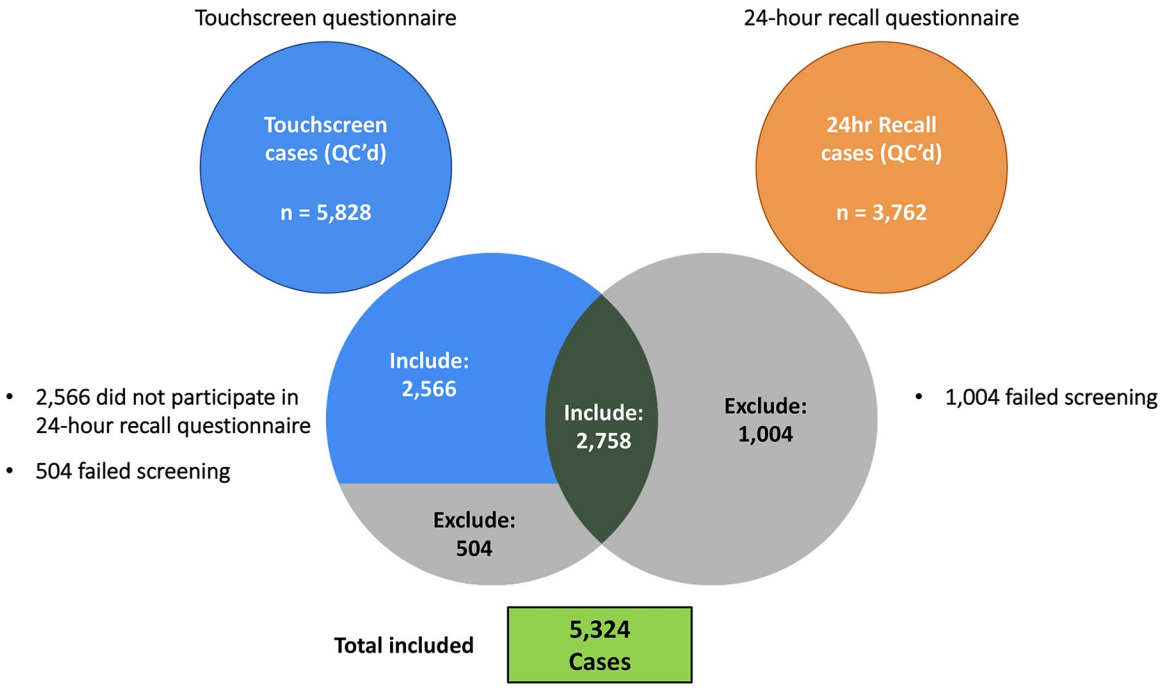

Fig 1. Case selection. Venn diagram showing vegetarian cases selected based on the Touchscreen and 24-hour recall questionnaires according to the criteria described in Methods section.

https://doi.org/10.1371/journal.pone.0291305.g001

controls) using Scalable and Accurate Implementation of GEneralized mixed model (SAIGE). SAIGE is particularly useful for imbalanced case-control ratios and has been previously tested on UKB data [56]. A detailed description of how the program is run can be found within the SAIGE github. The models were adjusted for genetic sex (data field 22001), age when attended the assessment center (data field 21003), assessment center attended for the initial assessment visit (data field 54), principal components 1 to 20 (data field 22009) and genotyping array and batch. The threshold for genome-wide significance was set at $P < 5 \times 10^{-8}$ and suggestive significance was set at $P < 5 \times 10^{-5}$. For further analysis of the GWAS data, we used the Functional Mapping and Annotation (FUMA) platform https://fuma.ctglab.nl/ [57, 58]. Its SNP2GENE function includes the Multi-marker Analysis of GenoMic Annotation (MAGMA) tool [59]. We used FUMA version 1.5.4 and MAGMA version 1.0.8.

Results

For the purposes of this study vegetarianism cases were defined as individuals who did not consume any animal flesh (including beef, lamb, pork, poultry, fish, and other seafood) or products derived from animal flesh, such as lard, for at least 1 year. Selection of vegetarianism cases was carried out using data from the touchscreen questionnaire, which provides diet information over the past year, and from the online 24-hour recall questionnaire, which focuses on food intake during the previous 24 hours. The touchscreen questionnaire was answered by approximately 500,000 respondents, whereas the 24-hour recall questionnaire had approximately 110,000 respondents. The touchscreen questionnaire was administered on 4 instances, and the 24-hour recall questionnaire was administered on 5 instances. A fraction of the participants responded to more than one instance of a questionnaire, and their responses for all instances were taken into consideration in determining cases and controls as

described in the Methods and Supporting Information sections in detail. Screening the two separate questionnaires resulted in two overlapping pools of individuals classed as vegetarianism cases (Fig 1). While the majority of these two pools overlapped, there were some individuals who failed one questionnaire screening but passed the other. These individuals were excluded from the total pool of cases. Some individuals who passed the touchscreen questionnaire were not included within the 24-hour recall cases, as not everyone who filled out the touchscreen questionnaire also filled out the 24-hour recall questionnaire. These individuals were retained for inclusion in the analysis. This resulted in a final pool of 5,324 cases (Fig 1). Any individual that indicated any level of intake of animal flesh based on the touchscreen questionnaire on at least one instance was retained for inclusion as a control. This resulted in a final pool of 329,455 controls.

Cases and controls were compared with regard to sex, age, body mass index, and Townsend deprivation index (a measure of socioeconomic status [60]) (Table 1). Vegetarianism cases differed significantly from controls on all 4 measures. Vegetarians were more likely to be women, younger in age, have a lower body mass index, and have a higher Townsend deprivation index (lower socioeconomic status). We then performed a logistic regression analysis using all 4 variables as predictors. They each remained highly significant in the multivariate analysis. These results are in agreement with previously reported data comparing vegetarians to non-vegetarians [61–66].

GWAS was performed for vegetarianism as a binary phenotype comparing controls to vegetarians and the results are summarized below.

There was some mild inflation as seen in the Q-Q plot of observed versus expected -$\log_{10}$ (*P*-value), with a genomic inflation factor (lambda) of 1.06 (95% CI 1.05–1.07) (Fig 2). Inflation due to population stratification was controlled in this GWAS by excluding related individuals, providing 20 principal components to account for cryptic relatedness and using SAIGE which accounts for sample relatedness. This slight inflation may therefore be due to the imbalance of cases and controls [56], population stratification which remains after correction from SAIGE, or another, unaccounted for confounding factor.

At the genome-wide significance threshold of 5×10^{-8}, one SNP on chromosome 18 was found to be significant: rs72884519, ($\beta = -0.11$, $P = 4.997 \times 10^{-8}$). A number of suggestively

Table 1. Characteristics of vegetarian and control populations.

	Total	Cases	Controls	p Value
N	334779	5324	329455	
Sex				<0.0001
Male	154585 (46%)	1788 (34%)	152797 (46%)	
Female	180194 (54%)	3536 (66%)	176658 (54%)	
Age (Mean ± SD)	56.9 ± 8.0	53.0 ± 7.9	56.9 ± 8.0	<0.0001
Male		53.0 ± 7.9	57.2 ± 8.1	
Female		53.0 ± 7.9	56.7 ± 7.9	
Townsend (Mean ± SD)	-1.60 ± 2.92	-1.01 ± 3.06	-1.61 ± 2.91	<0.0001
Male		-0.80 ± 3.20	-1.58 ± 2.97	
Female		-1.12 ± 2.98	-1.63 ± 2.87	
BMI (Mean ± SD)	27.4 ± 4.7	25.3 ± 4.5	27.4 ± 4.7	<0.0001
Male		25.7 ± 3.8	27.8 ± 4.2	
Female		25.1 ± 4.8	27.0 ± 5.1	

Vegetarian and control populations were compared for sex ratio by chi-square test, and for age, Townsend deprivation index, and body mass index (BMI) by t-test. The two populations were significantly different for all 4 variables. Logistic regression analysis was also performed using all 4 variables as predictors, and they each remained highly significant in the multivariate analysis.

https://doi.org/10.1371/journal.pone.0291305.t001

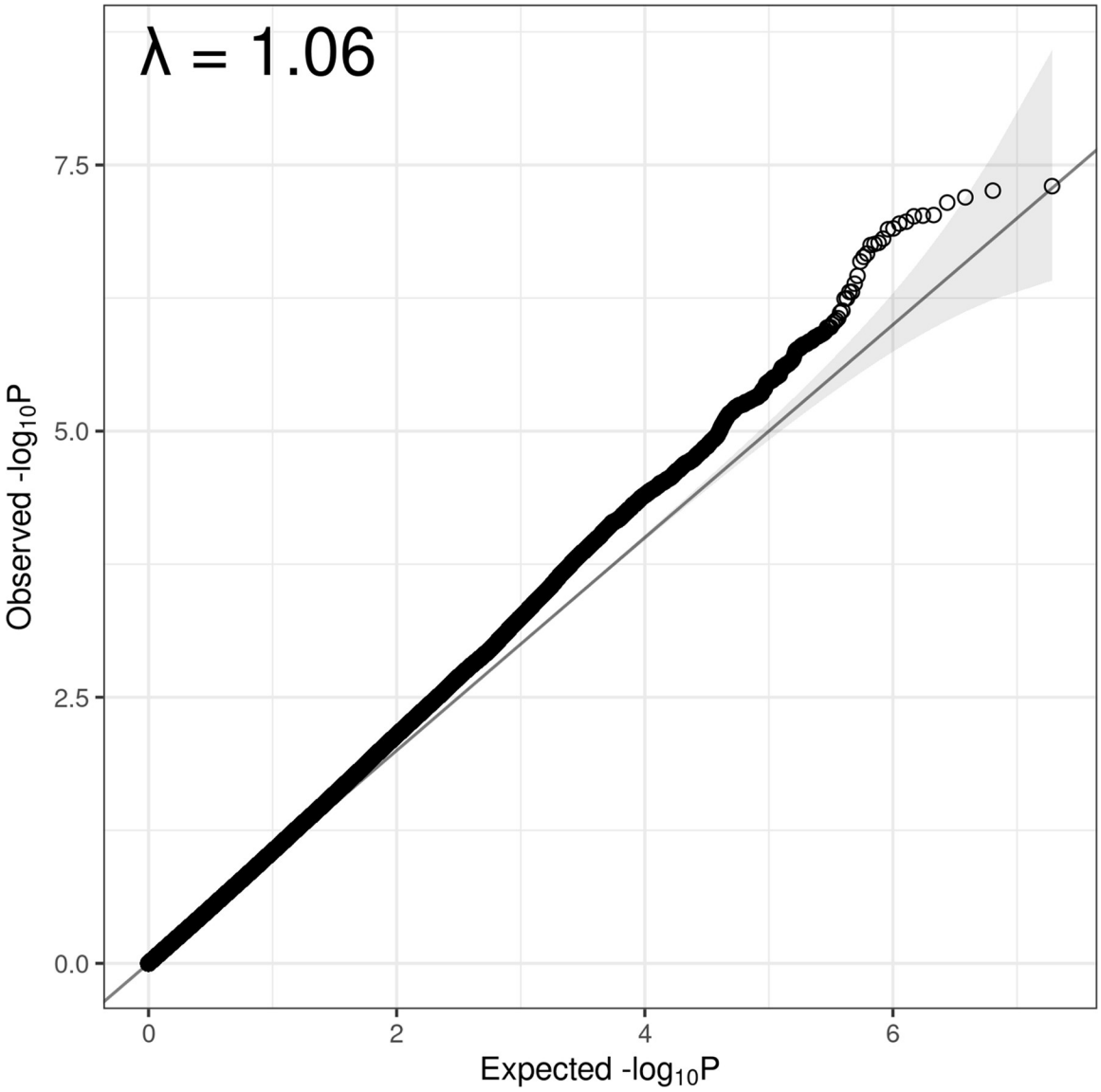

Fig 2. Q–Q plot. A Q–Q plot of observed versus expected -log10 (p-value). True associations are often expected for just a small minority of variants in the genome, in which case the Q–Q plot is expected to show a majority of points falling close to the expected 1:1 line, with a sharp upward curve in significance at the right-hand side of the plot (representing the true associations).

https://doi.org/10.1371/journal.pone.0291305.g002

significant SNPs in high linkage disequilibrium (LD) with rs72884519 were identified in the surrounding area, within genes *RIOK3*, *NPC1*, *RMC1* (C18orf8), and *TMEM241* (Figs 3 and 4, S1 Table).

Several other loci with groups of suggestively significant variants were identified, for a total of 202 significant and suggestively significant variants (Figs 3 and 4, S1 Table). An additional 7 genes associated with these loci contain at least one suggestively significant variant each (Fig 4 and S1 Table). This brings the total of genes that may contribute to the vegetarianism phenotype to 11: *RIOK3*, *NPC1*, *RMC1* (C18orf8), *TMEM241*, *VRK2*, *TMEM132D*, *METAP2*, *USP44*, *CDYL2*, *ZNF407*, *and CDH4*. A brief overview of what is known about the functions of these genes based on clinical and experimental data is provided in S5 Appendix, and a summary of traits associated with these genes identified by previous GWAS studies is provided in S2 Table.

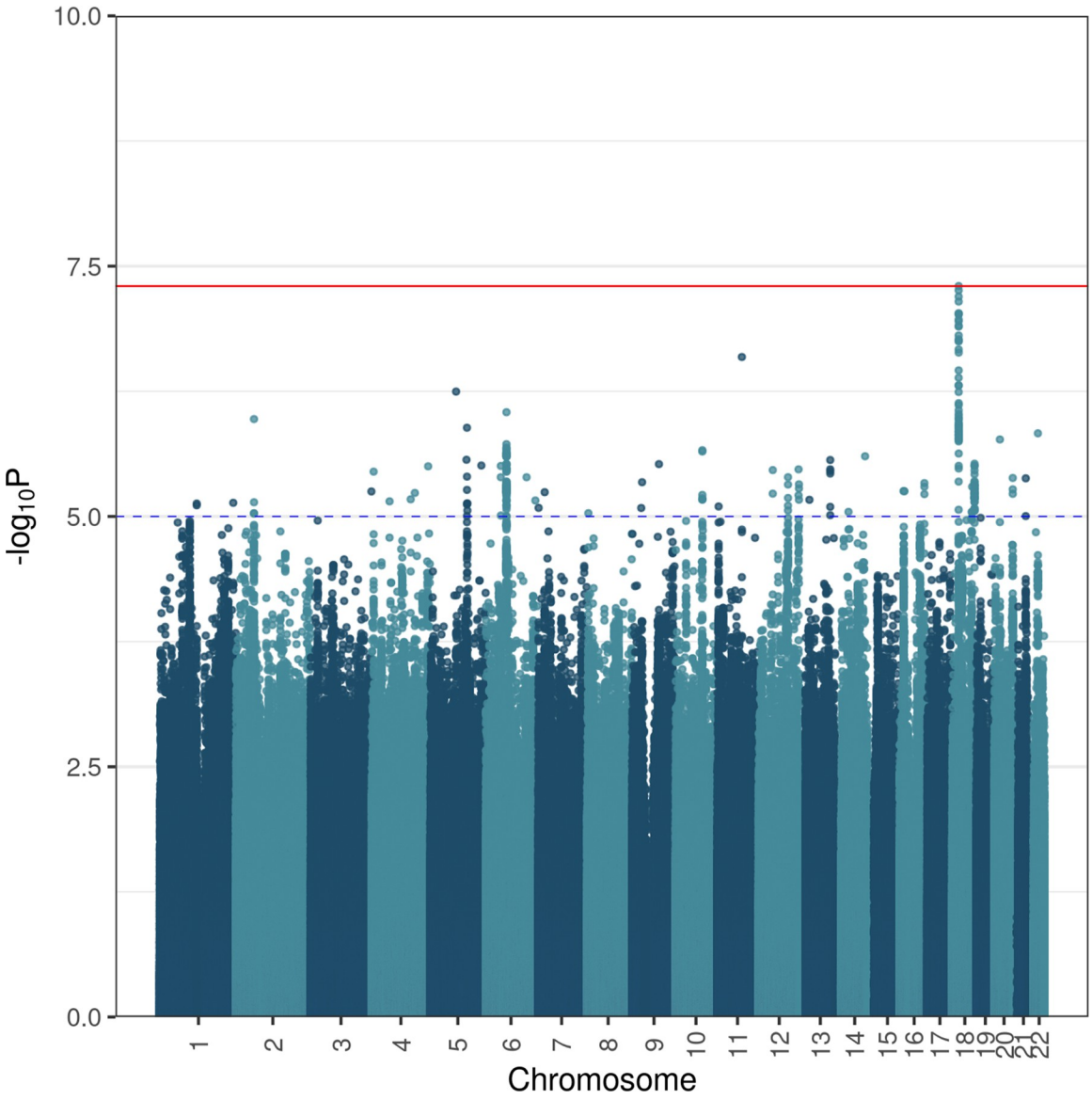

Fig 3. Manhattan plot. A Manhattan plot showing -log10 (p-value) for all variants as a function of position on the chromosomes. Sections of the genome showing association with vegetarianism cases can be identified as peaks within this plot. Due to linkage disequilibrium, adjacent variants are expected to show similar associations, and as a result, peaks of statistical significance may extend over several variants. The blue horizontal dashed line at -log10(1x10-5) represents the statistical threshold of suggestive association, while the red horizontal line at -log10(5x10-8) represents the statistical threshold of genome-wide significance.

https://doi.org/10.1371/journal.pone.0291305.g003

Further analysis of the GWAS data was carried out using the FUMA platform, which includes the MAGMA tool [57–59]. FUMA identified 37 genomic risk loci for vegetarianism and mapped 842 candidate SNPs and 59 genes to those loci (Fig 5 and S2–S4 Tables).

Gene-level GWAS analysis was carried out by MAGMA, which identified 3 genes with GWAS-significant p-values that clearly stand out from the rest: *RIOK3*, *RMC1*, and *NPC1* (Fig 6). The full MAGMA gene-level results are shown in S6 Table. Intersecting the mapped gene list with the MAGMA-generated gene list yields 34 genes with possible roles in vegetarianism, listed in S7 Table, sorted by MAGMA gene p-value.

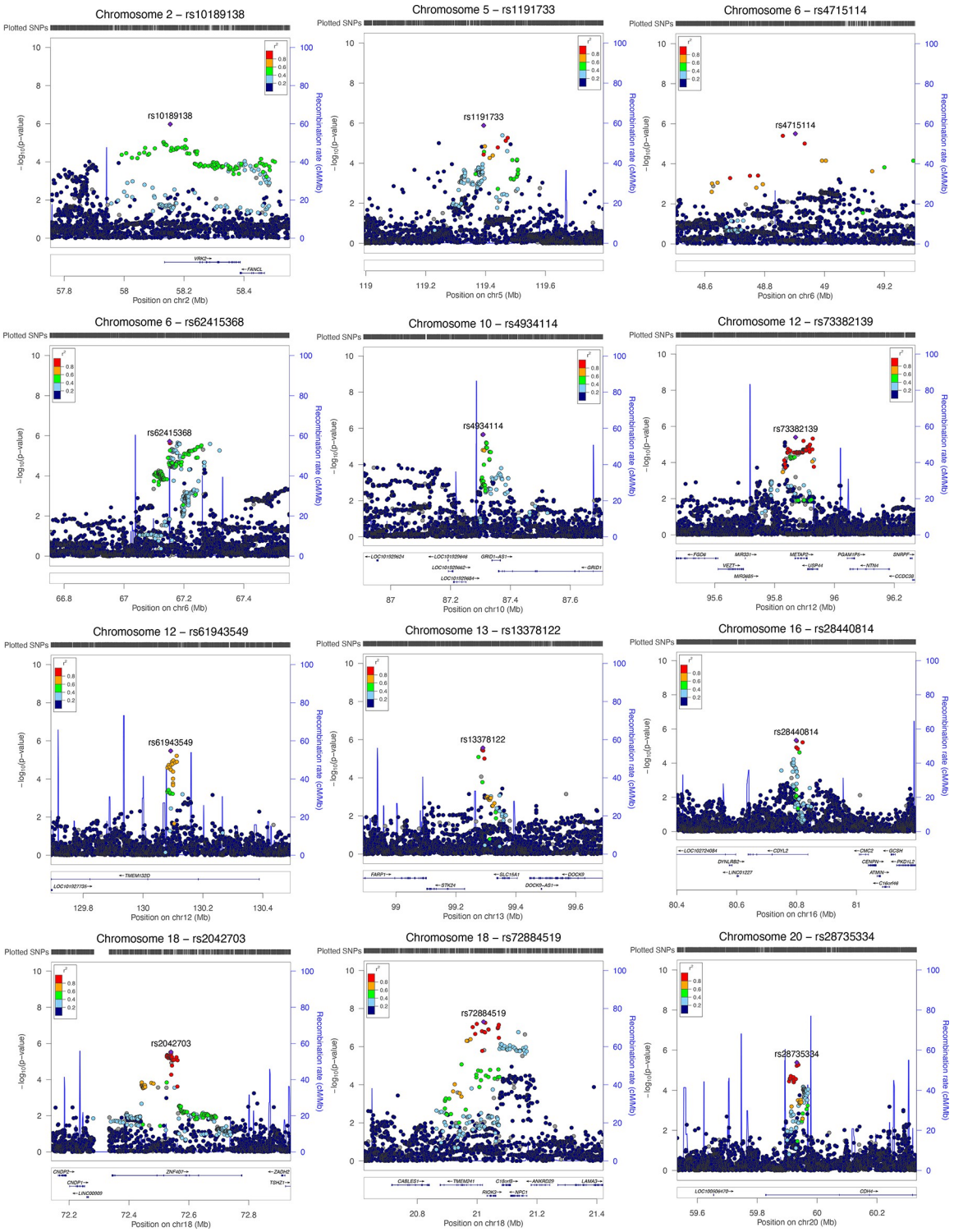

Fig 4. LocusZoom plots. LocusZoom plots are used to visualize GWAS results for 12 loci of interest. These loci include the one on chromosome 18 with the GWAS-significant SNP rs72884519, as well as loci with groups of 3 or more suggestively significant SNPs. LocusZoom plots display regional information including the strength and extent of the association signal relative to genomic position, local linkage disequilibrium (LD) and recombination patterns, and the positions of genes in the region. The x-axis shows the genomic position. The left-hand y-axis shows the -log10(p-value) as reported in S1 Table. The right-hand y-axis, which corresponds to the blue peaks within the plot, shows the recombination rate (cMMb) of that region of the chromosome. In each plot the variant with the lowest p-value in that region is shown as a purple diamond, while the other SNPs in the region are shown as small circles, color-coded to

indicate their degree of LD with that variant. Based on the LD score metric r2. An r2 of 0.8 would mean that the two SNPs are co-inherited roughly 80% of the time. In cases where there was no LD information available for a SNP (as is the case for chromosome 6 SNP rs62417319), the next most significant SNP has been plotted (rs62415368). Below the plot is the location of genes within the region, if present. These plots were generated using the original LocusZoom tool using the default settings. Summary statistics for the chromosome corresponding to the SNP of interest were used as input.

https://doi.org/10.1371/journal.pone.0291305.g004

Of the 3 GWAS-significant genes, *NPC1* is notable for being associated with SNPs that have the highest probability of being functionally relevant, with CADD score of 18.56 (rs1788799), RegulomeDB score of 1f (rs1624695), and commonChrState score of 1 (rs1623003) (S4 Table). The RegulomeDB parameter is used to predict the functionality of variants [67]; the CADD parameter prioritizes functional, deleterious and pathogenic variants [68]; and the commonChrState parameter indicates the most common 15-core chromatin state across 127 tissue/cell types [69]. *NPC1* (NPC Intracellular Cholesterol Transporter 1) encodes a large protein that resides in the membrane of endosomes and lysosomes and mediates intracellular trafficking of cholesterol and glycolipids [70–78]. Recent data indicate that activation of Rab7 by a trimeric GEF complex that includes RMC1, is required for lysosomal NPC1-dependent cholesterol export [79]. Mutations in *NPC1* are responsible for 95% of cases of Niemann-Pick disease type C, a lysosomal storage disease characterized by intracellular accumulation of cholesterol and glycosphingolipids in various tissues, with progressive neurological disease being the most significant clinical manifestation [80, 81]. As further discussed below, these findings

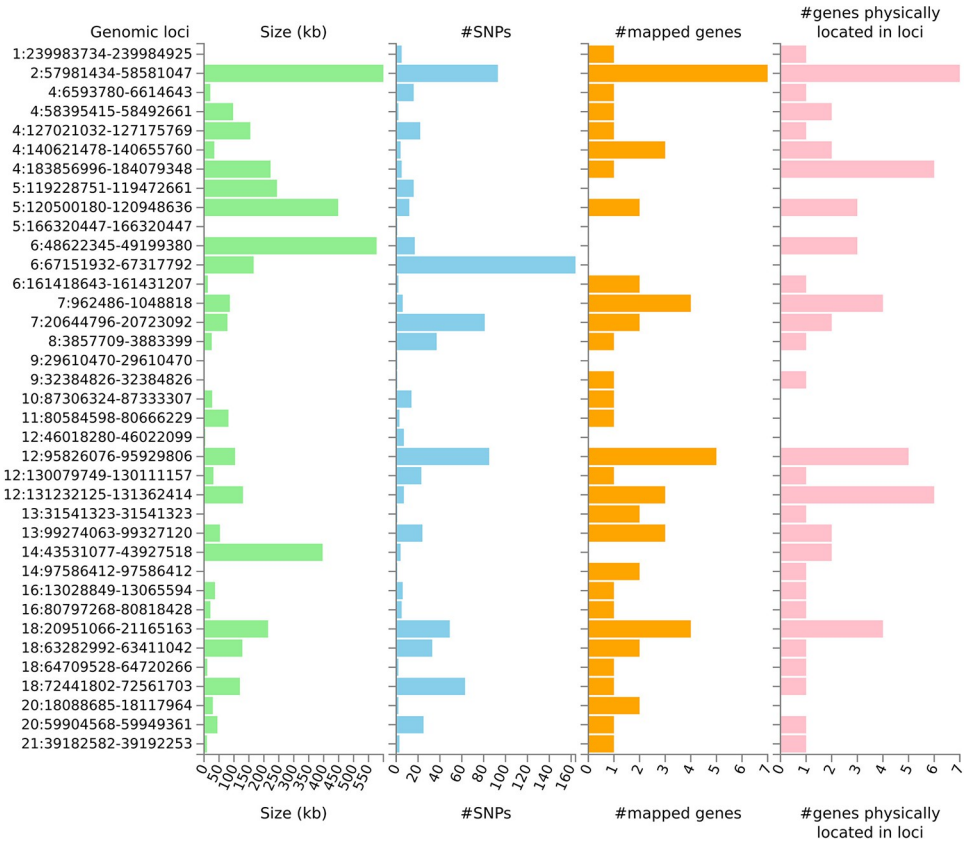

Fig 5. Summary of genomic risk loci. The histograms show summary results for each genomic locus identified by FUMA analysis. See text.

https://doi.org/10.1371/journal.pone.0291305.g005

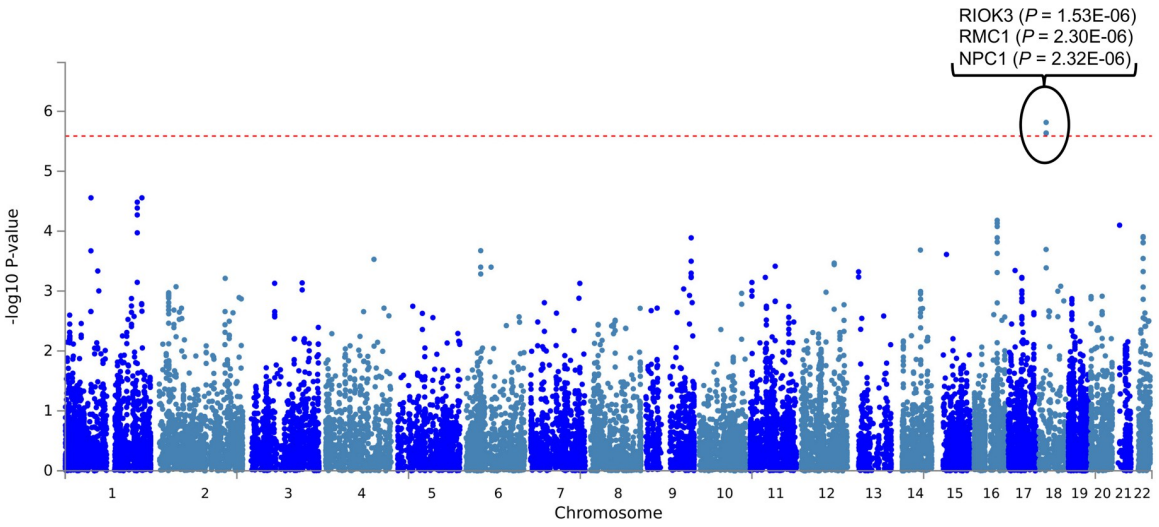

Fig 6. Gene-based Manhattan plot. Gene-based GWAS analysis was carried out using the MAGMA tool. The Manhattan plot shows -log10 (p-value) for all genes as a function of position on the chromosomes. Input SNPs were mapped to 19295 protein coding genes. Genome wide significance (red dashed line in the plot) was defined at P = 0.05/19295 = 2.591e-6.

https://doi.org/10.1371/journal.pone.0291305.g006

suggest that the genetic contribution to vegetarianism may be related to lipid metabolism and its role in brain function. This notion is supported by MAGMA tissue expression analysis showing that vegetarianism-associated variants may preferentially regulate gene expression in the brain (Fig 7). Additional support for this hypothesis is provided by FUMA analysis showing that SNPs associated with vegetarianism are also associated with several other traits in the GWAS Catalog that pertain to lipid metabolism and brain function (S8 Table).

Discussion

Numerous studies point to a significant genetic influence on food choice [34–48]. In this study we used GWAS to identify loci associated with long-term strict vegetarianism. Using data from UK Biobank, we identified one GWAS-significant SNP (rs72884519) and many suggestively significant variants (S1 Table, Figs and 4). Several of these suggestively significant variants are in high LD with rs72884519 (S1 Table, Figs 3 and 4), indicating that the result is not an artifact. The finding of only one GWAS-significant SNP could be due either to a lack of power, or lack of polygenicity. To investigate this, we performed a power calculation with CaTS (http://csg.sph.umich.edu//abecasis/CaTS). The power for a common SNP of frequency 50%, p-value of 10^{-7} and with a genotypic relative risk of 1.5 is 97% and 91% for a multiplicative and additive model, respectively, and 32% and 25% for a risk of 1.1. We chose a frequency of 50% since the power is much lower for rare SNPs, and a p-value of 10^{-7} to increase power. If there were hundreds of common causal SNPs, we would expect many more GWAS-significant ones even with modest power. We also performed LD score regression [82] and estimated the heritability to be 1.5%, with a 95% confidence interval of .9% - 15.2%. The intercept was estimated to be 1.01, indicating lack of population stratification. Overall, these data would support the hypothesis that our results may be due to a lack of polygenicity. More studies with larger sample sizes are needed to address this.

The significant and suggestively significant variants we identified are associated with 11 genes that may contribute to the vegetarianism phenotype. The functions of these genes are summarized in S5 Appendix, and their associations based on previous GWAS studies are

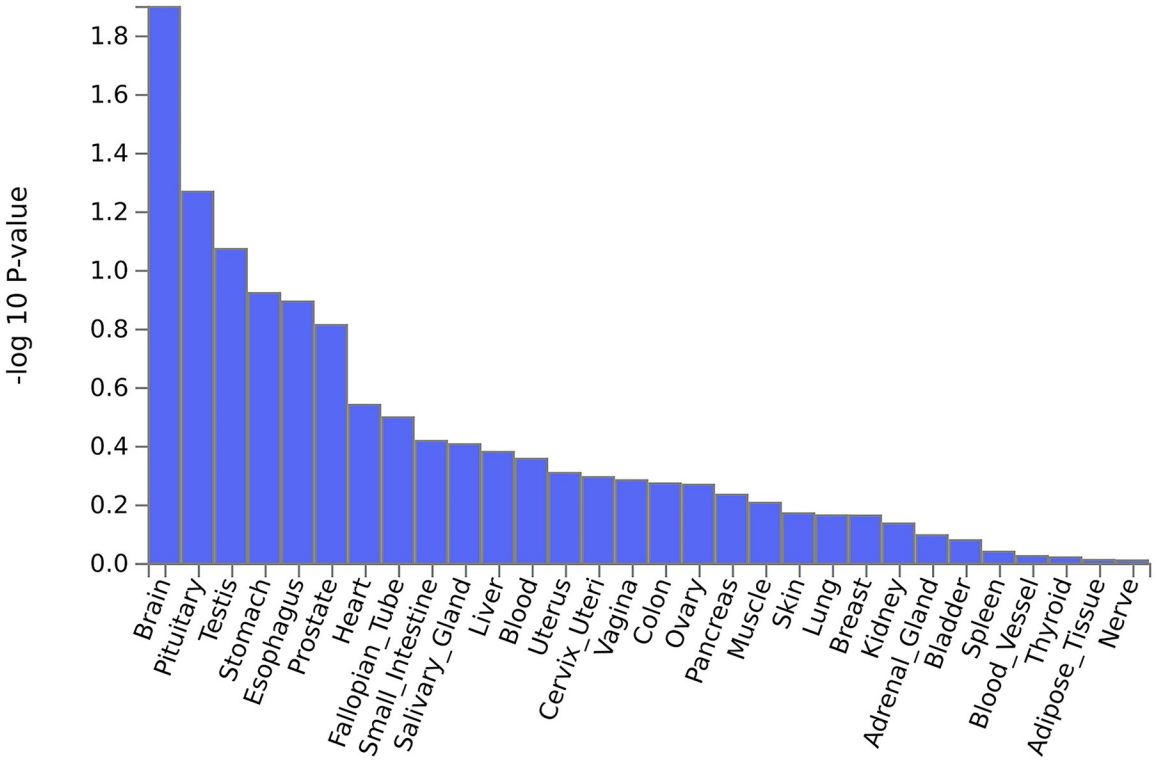

Fig 7. MAGMA tissue expression analysis: GTEx v8 30 general tissue types. To identify tissue specificity of the phenotype, FUMA performs MAGMA gene-property analyses to test relationships between tissue specific gene expression profiles and phenotype-gene associations. GTEx eQTL v8 contains 30 general tissue types.

https://doi.org/10.1371/journal.pone.0291305.g007

summarized in S2 Table. Those 11 genes were also among a set of 34 genes with possible roles in vegetarianism identified by MAGMA and FUMA analysis (S7 Table). Gene-level GWAS analysis by MAGMA showed that 3 of these genes, *RIOK3*, *RMC1*, and *NPC1*, are associated with vegetarianism at the GWAS-significant level (Fig 6 and S7 Table). In addition, previous GWAS studies have implicated some of these genes, including *RIOK3*, *NPC1*, *RMC1*, and *VRK2*, in dietary choices, lending further support to their role in vegetarianism (S2 Table).

The mechanisms by which genetic variants influence dietary choices involve an interplay between metabolism, physiologic effects, and taste perception. The levels of liking and consumption of dietary items are influenced by taste perception [83]. However, taste perception of a particular dietary item can be strongly influenced by its physiologic effects, which in turn are dictated by its metabolism. For example, while perception of bitterness impacts the levels of intake of caffeine [84, 85], there is evidence that the genetically determined rate of caffeine metabolism, which in turn determines its physiologic effects, influences both bitterness perception and the level of intake of caffeine, such that a low level of caffeine metabolism is associated with greater sensitivity to bitterness and a lower level of coffee consumption and vice versa [86–88]. Similarly, consumption of alcohol is strongly affected by genetic differences in enzymes involved in ethanol metabolism. Ethanol is converted to acetaldehyde by alcohol dehydrogenase (ADH1B) and acetaldehyde is further metabolized by aldehyde dehydrogenase 2 (ALDH2) to acetate. Since acetaldehyde has aversive effects, variants of ADH1B and ALDH2 that increase the levels of acetaldehyde are associated with aversion to alcohol, whereas variants associated with decreased acetaldehyde are associated with increased alcohol consumption and alcohol dependence [89].

It is therefore possible that choosing between a vegetarian and a non-vegetarian diet may also be dictated by individual differences in both metabolism and taste preferences. One study points to a genetic difference in lipid metabolism between vegetarians and non-vegetarians that appears to represent an adaptation to a vegetarian diet [90]. Long-chain polyunsaturated fatty acids (LCPUFA) are components of membrane phospholipids that play important roles in signal transduction; they are obtained from animal food sources or synthesized endogenously through a pathway that involves the fatty acid desaturases FADS1 and FADS2. An insertion in the *FADS2* gene that increases endogenous LCPUFA synthesis appears to have been selected for in vegetarian populations [90].

Interestingly, 2 of the 3 GWAS-significant genes associated with vegetarianism, *NPC1* and *RMC1*, also function in lipid metabolism. NPC1 mediates intracellular trafficking of cholesterol and glycolipids [70–78], and mutations in *NPC1* are responsible for 95% of cases of Niemann-Pick disease type C, a lysosomal storage disease characterized by intracellular accumulation of cholesterol and glycosphingolipids in various tissues [80, 81]. As discussed above, SNPs associated with *NPC1* have a high likelihood of being functionally relevant (S4 Table). RMC1 is necessary for cellular LDL-cholesterol uptake, and NPC1-dependent lysosomal cholesterol export [79, 91], and RMC1-deficient cells exhibit severe defects in LDL trafficking, with swelling of the late endosomal/lysosomal compartment and marked lysosomal cholesterol accumulation [79]. Another of the genes associated with vegetarianism, *TMEM241*, which has the 4th lowest p-value (S7 Table), is homologous to yeast *VRG4* that is involved in glycoprotein modification and mannosylation of sphingolipids [92], and decreased expression of *TMEM241* is associated with increased serum triglyceride levels in a Mexican population [93]. In addition, previous GWAS studies have shown that several vegetarianism-associated genes that we identified, including *TMEM241*, *NPC1*, *RMC1*, *RIOK3*, *VRK2*, and *TMEM132D*, are associated with markers of lipid metabolism and obesity such as triglyceride levels, LDL and HDL cholesterol levels, BMI, waist circumference, body fat, and others (S2 and S8 Tables).

Taken together, these data raise the possibility that differences in lipid metabolism may underlie the choice between a vegetarian and non-vegetarian diet. Lipid profiles of foods from animal sources are significantly different from those of plant sources, particularly with regard to complex lipids such as sphingolipids; and recent studies have brought attention to the importance of dietary sphingolipids in health [94–97]. In this regard it is of interest to note that complex lipids, particularly sphingolipids, play critical roles in nervous system development and function [98, 99], and that most of the vegetarianism-associated genes identified in this study have also been linked to psychological and neurological traits and diseases including Niemann-Pick disease, Alzheimer's disease, epilepsy, anxiety, depression, alcoholism, schizophrenia, autism spectrum disorder, cognitive performance and others (see S2 and S8 Tables). For example, *NPC1* mutations are responsible for the vast majority of cases of the neurological disorder Niemann-Pick disease type C [80, 81], and GWAS studies have shown links between *NPC1* and Alzheimer's disease [100]. Similarly, GWAS studies have linked *RIOK3*, *RMC1*, *NPC1*, *VRK2*, and *CDH4*, to educational attainment, cognitive performance, and/or alcohol consumption (S2 and S8 Tables). In addition, GWAS studies have shown associations between *VRK2* and size of the cerebral cortex, epilepsy, depression, bipolar disorder, anorexia nervosa, schizophrenia, autism spectrum disorder, and post-traumatic stress disorder (S2 Table). Of particular interest is the reported link between *VRK2* and anorexia nervosa [101]. Our data show that strict long-term vegetarianism is approximately twice as frequent in females (Table 1). Eating disorders are also more common in females and several studies suggest a link between vegetarianism and eating disorders [28, 102–108]. Thus, a better understanding of the metabolic underpinnings of vegetarianism may shed light on the mechanisms underlying

eating disorders, particularly anorexia nervosa. The roles that the genes we identified might play in neural function and in dietary choices remain to be determined. Based on the known functions and associations of some of these genes (S5 Appendix and, S2 and S8 Tables), one possibility is that the mechanisms involved may be related to the metabolism of complex lipids and its role in brain function.

Limitations of this study include the "healthy volunteer" selection bias in UK Biobank participants, who tend to be female, older, healthier, and of better socioeconomic status than the general population [109]. Our study was focused on one population of white UK persons in order to avoid confounding by ethnicity; it would be of interest for future studies to determine whether our findings can be replicated in other white and non-white populations. In addition, UK Biobank dietary data rely on self-reporting of prior food consumption. An interventional study would not be practical with a population of this size; however, once we have a better understanding of the gene variants associated with vegetarianism, it may become possible to carry out such a study prospectively to determine if a particular genetic signature can predict an individual's ability to adhere to a vegetarian diet. Finally, while our study is focused on the genetic contribution to strict long-term vegetarianism, dietary choices are not determined by genetics alone. Environmental factors and multiple disorders, and the medical interventions used to treat them, can impact dietary choices. Further studies are needed to determine the contributions of the genes we identified to the interplay between dietary choices and disorders of neural function and lipid metabolism.

In conclusion, our findings add to the existing body of data pointing to the genetic contribution to dietary choices and raise the possibility that lipid metabolism and its role in brain function may play a role in the ability to subsist long-term on a strict vegetarian diet. Further studies are required to determine which of the genes we identified play an important role in choosing a vegetarian diet; what particular variants of those genes underlie the vegetarianism phenotype and the mechanisms by which they contribute to this phenotype. It is tempting to speculate that meat may contain unique lipid components that vegetarians are able to adequately synthesize endogenously, whereas others need to obtain them from a meat-containing diet. A better understanding of the physiologic pathways involved in choosing a vegetarian diet is needed in order to design more effective dietary recommendations and interventions.

Supporting information

S1 Appendix. Sample QC.
(PDF)

S2 Appendix. SNP QC.
(PDF)

S3 Appendix. Phenotype processing methods.
(PDF)

S4 Appendix. GWAS methods.
(PDF)

S5 Appendix. Genes with possible roles in vegetarianism.
(PDF)

S1 Table. Vegetarianism-associated variants.
(XLSX)

S2 Table. Other genome-wide associations of vegetarianism-associated genes.
(DOCX)

S3 Table. FUMA genomic risk loci.
(XLSX)

S4 Table. FUMA candidate SNPs.
(XLSX)

S5 Table. FUMA mapped genes.
(XLSX)

S6 Table. MAGMA genes.
(XLSX)

S7 Table. FUMA-MAGMA genes intersect.
(XLSX)

S8 Table. FUMA GWAS catalog associations.
(XLSX)

Acknowledgments

Bioinformatics support was provided by Fios Genomics Ltd (Edinburgh, United Kingdom) as a commercial service.

Author Contributions

Conceptualization: Nabeel R. Yaseen, Akiko Takeda.

Data curation: Catriona L. K. Barnes, Lingwei Sun.

Formal analysis: Catriona L. K. Barnes, Lingwei Sun.

Funding acquisition: Nabeel R. Yaseen.

Investigation: Nabeel R. Yaseen.

Methodology: Catriona L. K. Barnes.

Project administration: Nabeel R. Yaseen.

Resources: Nabeel R. Yaseen.

Supervision: Nabeel R. Yaseen, John P. Rice.

Visualization: Nabeel R. Yaseen, Catriona L. K. Barnes.

Writing – original draft: Nabeel R. Yaseen.

Writing – review & editing: Catriona L. K. Barnes, Lingwei Sun, Akiko Takeda, John P. Rice.

References

1. Fox N, Ward K. Health, ethics and environment: a qualitative study of vegetarian motivations. Appetite. 2008; 50(2–3):422–9. Epub 2007/11/06. https://doi.org/10.1016/j.appet.2007.09.007 PMID: 17980457.
2. Ruby MB. Vegetarianism. A blossoming field of study. Appetite. 2012; 58(1):141–50. Epub 2011/10/18. https://doi.org/10.1016/j.appet.2011.09.019 PMID: 22001025.

3. Bhattacharya M. A historical exploration of Indian diets and a possible link to insulin resistance syndrome. Appetite. 2015; 95:421–54. Epub 2015/07/25. https://doi.org/10.1016/j.appet.2015.07.002 PMID: 26206172.

4. Leitzmann C. Vegetarian nutrition: past, present, future. Am J Clin Nutr. 2014; 100 Suppl 1:496S–502S. Epub 2014/06/06. https://doi.org/10.3945/ajcn.113.071365 PMID: 24898226.

5. Craig WJ, Mangels AR. Position of the American Dietetic Association: vegetarian diets. J Am Diet Assoc. 2009; 109(7):1266–82. Epub 2009/07/01. https://doi.org/10.1016/j.jada.2009.05.027 PMID: 19562864.

6. Position of the American Dietetic Association and Dietitians of Canada: vegetarian diets. Can J Diet Pract Res. 2003; 64(2):62–81. Epub 2003/06/27. https://doi.org/10.3148/64.2.2003.62 PMID: 12826028.

7. Lea EJ, Crawford D, Worsley A. Public views of the benefits and barriers to the consumption of a plant-based diet. Eur J Clin Nutr. 2006; 60(7):828–37. Epub 2006/02/03. https://doi.org/10.1038/sj.ejcn.1602387 PMID: 16452915.

8. Battaglia Richi E, Baumer B, Conrad B, Darioli R, Schmid A, Keller U. Health Risks Associated with Meat Consumption: A Review of Epidemiological Studies. Int J Vitam Nutr Res. 2015; 85(1–2):70–8. Epub 2016/01/19. https://doi.org/10.1024/0300-9831/a000224 PMID: 26780279.

9. Dybvik JS, Svendsen M, Aune D. Vegetarian and vegan diets and the risk of cardiovascular disease, ischemic heart disease and stroke: a systematic review and meta-analysis of prospective cohort studies. Eur J Nutr. 2022. Epub 20220827. https://doi.org/10.1007/s00394-022-02942-8 PMID: 36030329.

10. Key TJ, Appleby PN, Spencer EA, Travis RC, Allen NE, Thorogood M, et al. Cancer incidence in British vegetarians. Br J Cancer. 2009; 101(1):192–7. Epub 20090616. https://doi.org/10.1038/sj.bjc.6605098 PMID: 19536095; PubMed Central PMCID: PMC2699384.

11. Rizzo NS, Sabaté J, Jaceldo-Siegl K, Fraser GE. Vegetarian dietary patterns are associated with a lower risk of metabolic syndrome: the adventist health study 2. Diabetes Care. 2011; 34(5):1225–7. Epub 20110316. https://doi.org/10.2337/dc10-1221 PMID: 21411506; PubMed Central PMCID: PMC3114510.

12. Jarvis SE, Nguyen M, Malik V. Association between adherence to plant-based dietary patterns and obesity risk: a systematic review of prospective cohort studies. Appl Physiol Nutr Metab. 2022. Epub 20220819. https://doi.org/10.1139/apnm-2022-0059 PMID: 35985038.

13. Dawczynski C, Weidauer T, Richert C, Schlattmann P, Dawczynski K, Kiehntopf M. Nutrient Intake and Nutrition Status in Vegetarians and Vegans in Comparison to Omnivores—the Nutritional Evaluation (NuEva) Study. Front Nutr. 2022; 9:819106. Epub 20220516. https://doi.org/10.3389/fnut.2022.819106 PMID: 35651513; PubMed Central PMCID: PMC9149309.

14. Key TJ, Papier K, Tong TYN. Plant-based diets and long-term health: findings from the EPIC-Oxford study. Proc Nutr Soc. 2022; 81(2):190–8. Epub 20211027. https://doi.org/10.1017/S0029665121003748 PMID: 35934687.

15. DeClercq V, Nearing JT, Sweeney E. Plant-Based Diets and Cancer Risk: What is the Evidence? Curr Nutr Rep. 2022; 11(2):354–69. Epub 20220325. https://doi.org/10.1007/s13668-022-00409-0 PMID: 35334103.

16. Fan SM, Chiu PY, Liu CH, Liao YC, Chang HT. Predictive value of hypercholesterolemia, vegetarian diet, and hypertension for incident dementia among elderly Taiwanese individuals with low educational levels. Ther Adv Chronic Dis. 2023; 14:20406223231171549. Epub 20230527. https://doi.org/10.1177/20406223231171549 PMID: 37255548; PubMed Central PMCID: PMC10226334.

17. Jack ML, Sumrall K, Nasrallah C, Stuckey D, Jotwani V. Analysis of Serum Ferritin Levels in a Group of Elite Ballet Dancers. J Dance Med Sci. 2023:1089313x231178089. Epub 20230601. https://doi.org/10.1177/1089313x231178089 PMID: 37264629.

18. Jain R, Larsuphrom P, Degremont A, Latunde-Dada GO, Philippou E. Association between vegetarian and vegan diets and depression: A systematic review. Nutr Bull. 2022; 47(1):27–49. Epub 20220119. https://doi.org/10.1111/nbu.12540 PMID: 36045075.

19. Smits KPJ, Listl S, Jevdjevic M. Vegetarian diet and its possible influence on dental health: A systematic literature review. Community Dent Oral Epidemiol. 2020; 48(1):7–13. Epub 20191001. https://doi.org/10.1111/cdoe.12498 PMID: 31571246; PubMed Central PMCID: PMC6972589.

20. Tsai CK, Nfor ON, Tantoh DM, Lu WY, Liaw YP. The association between vegetarian diet and varicose veins might be more prominent in men than in women. Front Nutr. 2023; 10:1046158. Epub 20230601. https://doi.org/10.3389/fnut.2023.1046158 PMID: 37324727; PubMed Central PMCID: PMC10267867.

21. Zheng Y, Wang J, Wang Y, Xu K, Chen X. The Hidden Dangers of Plant-Based Diets Affecting Bone Health: A Cross-Sectional Study with U.S. National Health and Nutrition Examination Survey

(NHANES) Data from 2005–2018. Nutrients. 2023; 15(7). Epub 20230406. https://doi.org/10.3390/nu15071794 PMID: 37049634; PubMed Central PMCID: PMC10097387.

22. Bai S, Zhang Y, Guo C, Liu Y, Zhang Q, Liu L, et al. Associations between dietary patterns and nephrolithiasis risk in a large Chinese cohort: is a balanced or plant-based diet better? Food Funct. 2023; 14 (7):3220–9. Epub 20230403. https://doi.org/10.1039/d2fo03993a PMID: 36920109.

23. Bali A, Naik R. The Impact of a Vegan Diet on Many Aspects of Health: The Overlooked Side of Veganism. Cureus. 2023; 15(2):e35148. Epub 20230218. https://doi.org/10.7759/cureus.35148 PMID: 36950003; PubMed Central PMCID: PMC10027313.

24. Chirravuri V, Ghonge S, Palal D. Cross-Sectional Study of Serum Vitamin B12 and Vitamin D3 Levels Amongst Corporate Employees. Cureus. 2023; 15(2):e34642. Epub 20230205. https://doi.org/10.7759/cureus.34642 PMID: 36751576; PubMed Central PMCID: PMC9899440.

25. Eveleigh ER, Coneyworth L, Welham SJM. Systematic review and meta-analysis of iodine nutrition in modern vegan and vegetarian diets. Br J Nutr. 2023:1–15. Epub 20230313. https://doi.org/10.1017/S000711452300051X PMID: 36912094.

26. Sinclair AJ, Wang Y, Li D. What Is the Evidence for Dietary-Induced DHA Deficiency in Human Brains? Nutrients. 2022; 15(1). Epub 20221229. https://doi.org/10.3390/nu15010161 PMID: 36615819; PubMed Central PMCID: PMC9824463.

27. Weder S, Zerback EH, Wagener SM, Koeder C, Fischer M, Alexy U, et al. How Does Selenium Intake Differ among Children (1–3 Years) on Vegetarian, Vegan, and Omnivorous Diets? Results of the VeChi Diet Study. Nutrients. 2022; 15(1). Epub 20221221. https://doi.org/10.3390/nu15010034 PMID: 36615692; PubMed Central PMCID: PMC9824336.

28. Robinson-O'Brien R, Perry CL, Wall MM, Story M, Neumark-Sztainer D. Adolescent and young adult vegetarianism: better dietary intake and weight outcomes but increased risk of disordered eating behaviors. J Am Diet Assoc. 2009; 109(4):648–55. https://doi.org/10.1016/j.jada.2008.12.014 PMID: 19328260.

29. You W, Henneberg R, Saniotis A, Ge Y, Henneberg M. Total Meat Intake is Associated with Life Expectancy: A Cross-Sectional Data Analysis of 175 Contemporary Populations. Int J Gen Med. 2022; 15:1833–51. Epub 20220222. https://doi.org/10.2147/IJGM.S333004 PMID: 35228814; PubMed Central PMCID: PMC8881926.

30. Stahler C. How many adults are vegetarian? Veg J. 2006; 25:14–5.

31. Barr SI, Chapman GE. Perceptions and practices of self-defined current vegetarian, former vegetarian, and nonvegetarian women. J Am Diet Assoc. 2002; 102(3):354–60. Epub 2002/03/21. https://doi.org/10.1016/s0002-8223(02)90083-0 PMID: 11902368.

32. Haddad EH, Tanzman JS. What do vegetarians in the United States eat? Am J Clin Nutr. 2003; 78(3 Suppl):626S–32S. Epub 2003/08/26. https://doi.org/10.1093/ajcn/78.3.626S PMID: 12936957.

33. Juan WY, Yamini S, Britten P. Food Intake Patterns of Self-identified Vegetarians among the U.S. Population, 2007–2010. Procedia Food Science. 2015; 4:86–93.

34. Breen FM, Plomin R, Wardle J. Heritability of food preferences in young children. Physiol Behav. 2006; 88(4–5):443–7. Epub 2006/06/06. https://doi.org/10.1016/j.physbeh.2006.04.016 [pii] PMID: 16750228.

35. de Castro JM. Genetic influences on daily intake and meal patterns of humans. Physiol Behav. 1993; 53(4):777–82. Epub 1993/04/01. https://doi.org/10.1016/0031-9384(93)90188-l [pii]. PMID: 8511185.

36. Gunderson EP, Tsai AL, Selby JV, Caan B, Mayer-Davis EJ, Risch N. Twins of mistaken zygosity (TOMZ): evidence for genetic contributions to dietary patterns and physiologic traits. Twin Res Hum Genet. 2006; 9(4):540–9. Epub 2006/08/11. https://doi.org/10.1375/183242706778025053 PMID: 16899161.

37. Heitmann BL, Harris JR, Lissner L, Pedersen NL. Genetic effects on weight change and food intake in Swedish adult twins. Am J Clin Nutr. 1999; 69(4):597–602. Epub 1999/04/10. https://doi.org/10.1093/ajcn/69.4.597 PMID: 10197559.

38. Heller RF, O'Connell DL, Roberts DC, Allen JR, Knapp JC, Steele PL, et al. Lifestyle factors in monozygotic and dizygotic twins. Genet Epidemiol. 1988; 5(5):311–21. Epub 1988/01/01. https://doi.org/10.1002/gepi.1370050503 PMID: 3215506.

39. Hur YM, Bouchard TJ Jr., Eckert E. Genetic and environmental influences on self-reported diet: a reared-apart twin study. Physiol Behav. 1998; 64(5):629–36. Epub 1998/11/17. https://doi.org/10.1016/s0031-9384(98)00101-2 PMID: 9817574.

40. Kaprio J, Pulkkinen L, Rose RJ. Genetic and environmental factors in health-related behaviors: studies on Finnish twins and twin families. Twin Res. 2002; 5(5):366–71. Epub 2003/01/23. https://doi.org/10.1375/136905202320906101 PMID: 12537860.

41. Keller KL, Pietrobelli A, Must S, Faith MS. Genetics of eating and its relation to obesity. Curr Atheroscler Rep. 2002; 4(3):176–82. Epub 2002/04/05. https://doi.org/10.1007/s11883-002-0017-3 PMID: 11931714.

42. Keskitalo K, Silventoinen K, Tuorila H, Perola M, Pietilainen KH, Rissanen A, et al. Genetic and environmental contributions to food use patterns of young adult twins. Physiol Behav. 2008; 93(1–2):235–42. Epub 2007/09/28. https://doi.org/10.1016/j.physbeh.2007.08.025 [pii] PMID: 17897688.

43. Krondl M, Coleman P, Wade J, Milner J. A twin study examining the genetic influence on food selection. Hum Nutr Appl Nutr. 1983; 37 A(3):189–98. Epub 1983/06/01. PMID: 6683717.

44. Teucher B, Skinner J, Skidmore PM, Cassidy A, Fairweather-Tait SJ, Hooper L, et al. Dietary patterns and heritability of food choice in a UK female twin cohort. Twin Res Hum Genet. 2007; 10(5):734–48. Epub 2007/10/02. https://doi.org/10.1375/twin.10.5.734 PMID: 17903115.

45. van den Bree MB, Eaves LJ, Dwyer JT. Genetic and environmental influences on eating patterns of twins aged >/ = 50 y. Am J Clin Nutr. 1999; 70(4):456–65. Epub 1999/09/29. https://doi.org/10.1093/ajcn/70.4.456 PMID: 10500013.

46. Smith AD, Fildes A, Cooke L, Herle M, Shakeshaft N, Plomin R, et al. Genetic and environmental influences on food preferences in adolescence. Am J Clin Nutr. 2016; 104(2):446–53. Epub 2016/07/08. https://doi.org/10.3945/ajcn.116.133983 PMID: 27385609; PubMed Central PMCID: PMC4962164.

47. May-Wilson S, Matoba N, Wade KH, Hottenga JJ, Concas MP, Mangino M, et al. Large-scale GWAS of food liking reveals genetic determinants and genetic correlations with distinct neurophysiological traits. Nat Commun. 2022; 13(1):2743. Epub 2022/05/19. https://doi.org/10.1038/s41467-022-30187-w PMID: 35585065; PubMed Central PMCID: PMC9117208.

48. Vink JM, van Hooijdonk KJM, Willemsen G, Feskens EJM, Boomsma DI. Causes of Variation in Food Preference in the Netherlands. Twin Res Hum Genet. 2020; 23(4):195–203. Epub 2020/09/05. https://doi.org/10.1017/thg.2020.66 PMID: 32885771.

49. Mompeo O, Gibson R, Christofidou P, Spector TD, Menni C, Mangino M. Genetic and Environmental Influences of Dietary Indices in a UK Female Twin Cohort. Twin Res Hum Genet. 2020; 23(6):330–7. Epub 20210118. https://doi.org/10.1017/thg.2020.84 PMID: 33455612.

50. Pallister T, Sharafi M, Lachance G, Pirastu N, Mohney RP, MacGregor A, et al. Food Preference Patterns in a UK Twin Cohort. Twin Res Hum Genet. 2015; 18(6):793–805. Epub 20150928. https://doi.org/10.1017/thg.2015.69 PMID: 26412323.

51. Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, et al. The UK Biobank resource with deep phenotyping and genomic data. Nature. 2018; 562(7726):203–9. https://doi.org/10.1038/s41586-018-0579-z PMID: 30305743

52. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, et al. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. PLoS Med. 2015; 12(3):e1001779. Epub 20150331. https://doi.org/10.1371/journal.pmed.1001779 PMID: 25826379; PubMed Central PMCID: PMC4380465.

53. Howie B, Fuchsberger C, Stephens M, Marchini J, Abecasis GR. Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. Nature Genetics. 2012; 44(8):955–9. https://doi.org/10.1038/ng.2354 PMID: 22820512

54. McCarthy S, Das S, Kretzschmar W, Delaneau O, Wood AR, Teumer A, et al. A reference panel of 64,976 haplotypes for genotype imputation. Nat Genet. 2016; 48(10):1279–83. Epub 20160822. https://doi.org/10.1038/ng.3643 PMID: 27548312; PubMed Central PMCID: PMC5388176.

55. Huang J, Howie B, McCarthy S, Memari Y, Walter K, Min JL, et al. Improved imputation of low-frequency and rare variants using the UK10K haplotype reference panel. Nat Commun. 2015; 6:8111. Epub 20150914. https://doi.org/10.1038/ncomms9111 PMID: 26368830; PubMed Central PMCID: PMC4579394.

56. Zhou W, Nielsen JB, Fritsche LG, Dey R, Gabrielsen ME, Wolford BN, et al. Efficiently controlling for case-control imbalance and sample relatedness in large-scale genetic association studies. Nature Genetics. 2018; 50(9):1335–41. https://doi.org/10.1038/s41588-018-0184-y PMID: 30104761

57. Watanabe K, Taskesen E, van Bochoven A, Posthuma D. Functional mapping and annotation of genetic associations with FUMA. Nature Communications. 2017; 8(1):1826. https://doi.org/10.1038/s41467-017-01261-5 PMID: 29184056

58. Watanabe K, Umićević Mirkov M, de Leeuw CA, van den Heuvel MP, Posthuma D. Genetic mapping of cell type specificity for complex traits. Nature Communications. 2019; 10(1):3222. https://doi.org/10.1038/s41467-019-11181-1 PMID: 31324783

59. de Leeuw CA, Mooij JM, Heskes T, Posthuma D. MAGMA: generalized gene-set analysis of GWAS data. PLoS Comput Biol. 2015; 11(4):e1004219. Epub 20150417. https://doi.org/10.1371/journal.pcbi.1004219 PMID: 25885710; PubMed Central PMCID: PMC4401657.

60. Townsend P, Phillimore P, Beattie A. Health and Deprivation: Inequality and the North. London: Croom Helm; 1988.

61. Tong TY, Key TJ, Sobiecki JG, Bradbury KE. Anthropometric and physiologic characteristics in white and British Indian vegetarians and nonvegetarians in the UK Biobank. Am J Clin Nutr. 2018; 107 (6):909–20. https://doi.org/10.1093/ajcn/nqy042 PMID: 29868910; PubMed Central PMCID: PMC5985736.

62. Tong TYN, Perez-Cornago A, Bradbury KE, Key TJ. Biomarker Concentrations in White and British Indian Vegetarians and Nonvegetarians in the UK Biobank. J Nutr. 2021; 151(10):3168–79. https://doi.org/10.1093/jn/nxab192 PMID: 34132352; PubMed Central PMCID: PMC8485916.

63. Spencer EA, Appleby PN, Davey GK, Key TJ. Diet and body mass index in 38 000 EPIC-Oxford meat-eaters, fish-eaters, vegetarians and vegans. International Journal of Obesity. 2003; 27(6):728–34. https://doi.org/10.1038/sj.ijo.0802300 PMID: 12833118

64. Tonstad S, Butler T, Yan R, Fraser GE. Type of vegetarian diet, body weight, and prevalence of type 2 diabetes. Diabetes Care. 2009; 32(5):791–6. Epub 20090407. https://doi.org/10.2337/dc08-1886 PMID: 19351712; PubMed Central PMCID: PMC2671114.

65. Tong TYN, Appleby PN, Armstrong MEG, Fensom GK, Knuppel A, Papier K, et al. Vegetarian and vegan diets and risks of total and site-specific fractures: results from the prospective EPIC-Oxford study. BMC Med. 2020; 18(1):353. Epub 20201123. https://doi.org/10.1186/s12916-020-01815-3 PMID: 33222682; PubMed Central PMCID: PMC7682057.

66. Crowe FL, Appleby PN, Allen NE, Key TJ. Diet and risk of diverticular disease in Oxford cohort of European Prospective Investigation into Cancer and Nutrition (EPIC): prospective study of British vegetarians and non-vegetarians. Bmj. 2011; 343:d4131. Epub 20110719. https://doi.org/10.1136/bmj.d4131 PMID: 21771850; PubMed Central PMCID: PMC3139912.

67. Boyle AP, Hong EL, Hariharan M, Cheng Y, Schaub MA, Kasowski M, et al. Annotation of functional variation in personal genomes using RegulomeDB. Genome Res. 2012; 22(9):1790–7. https://doi.org/10.1101/gr.137323.112 PMID: 22955989; PubMed Central PMCID: PMC3431494.

68. Kircher M, Witten DM, Jain P, O'Roak BJ, Cooper GM, Shendure J. A general framework for estimating the relative pathogenicity of human genetic variants. Nature Genetics. 2014; 46(3):310–5. https://doi.org/10.1038/ng.2892 PMID: 24487276

69. Kundaje A, Meuleman W, Ernst J, Bilenky M, Yen A, Heravi-Moussavi A, et al. Integrative analysis of 111 reference human epigenomes. Nature. 2015; 518(7539):317–30. https://doi.org/10.1038/nature14248 PMID: 25693563; PubMed Central PMCID: PMC4530010.

70. Watari H, Blanchette-Mackie EJ, Dwyer NK, Glick JM, Patel S, Neufeld EB, et al. Niemann-Pick C1 protein: obligatory roles for N-terminal domains and lysosomal targeting in cholesterol mobilization. Proc Natl Acad Sci U S A. 1999; 96(3):805–10. Epub 1999/02/03. https://doi.org/10.1073/pnas.96.3.805 PMID: 9927649; PubMed Central PMCID: PMC15306.

71. Carstea ED, Morris JA, Coleman KG, Loftus SK, Zhang D, Cummings C, et al. Niemann-Pick C1 disease gene: homology to mediators of cholesterol homeostasis. Science. 1997; 277(5323):228–31. https://doi.org/10.1126/science.277.5323.228 PMID: 9211849.

72. Davies JP, Ioannou YA. Topological analysis of Niemann-Pick C1 protein reveals that the membrane orientation of the putative sterol-sensing domain is identical to those of 3-hydroxy-3-methylglutaryl-CoA reductase and sterol regulatory element binding protein cleavage-activating protein. J Biol Chem. 2000; 275(32):24367–74. https://doi.org/10.1074/jbc.M002184200 PMID: 10821832.

73. Infante RE, Wang ML, Radhakrishnan A, Kwon HJ, Brown MS, Goldstein JL. NPC2 facilitates bidirectional transfer of cholesterol between NPC1 and lipid bilayers, a step in cholesterol egress from lysosomes. Proc Natl Acad Sci U S A. 2008; 105(40):15287–92. Epub 20080904. https://doi.org/10.1073/pnas.0807328105 PMID: 18772377; PubMed Central PMCID: PMC2563079.

74. Gong X, Qian H, Zhou X, Wu J, Wan T, Cao P, et al. Structural Insights into the Niemann-Pick C1 (NPC1)-Mediated Cholesterol Transfer and Ebola Infection. Cell. 2016; 165(6):1467–78. Epub 20160526. https://doi.org/10.1016/j.cell.2016.05.022 PMID: 27238017; PubMed Central PMCID: PMC7111323.

75. Blom TS, Linder MD, Snow K, Pihko H, Hess MW, Jokitalo E, et al. Defective endocytic trafficking of NPC1 and NPC2 underlying infantile Niemann-Pick type C disease. Hum Mol Genet. 2003; 12 (3):257–72. https://doi.org/10.1093/hmg/ddg025 PMID: 12554680.

76. Kwon HJ, Abi-Mosleh L, Wang ML, Deisenhofer J, Goldstein JL, Brown MS, et al. Structure of N-terminal domain of NPC1 reveals distinct subdomains for binding and transfer of cholesterol. Cell. 2009; 137(7):1213–24. https://doi.org/10.1016/j.cell.2009.03.049 PMID: 19563754; PubMed Central PMCID: PMC2739658.

77. Li X, Lu F, Trinh MN, Schmiege P, Seemann J, Wang J, et al. 3.3 A structure of Niemann-Pick C1 protein reveals insights into the function of the C-terminal luminal domain in cholesterol transport. Proc

Natl Acad Sci U S A. 2017; 114(34):9116–21. Epub 20170807. https://doi.org/10.1073/pnas.1711716114 PMID: 28784760; PubMed Central PMCID: PMC5576846.

78. Zhang M, Dwyer NK, Neufeld EB, Love DC, Cooney A, Comly M, et al. Sterol-modulated glycolipid sorting occurs in niemann-pick C1 late endosomes. J Biol Chem. 2001; 276(5):3417–25. Epub 20001013. https://doi.org/10.1074/jbc.M005393200 PMID: 11032830.

79. van den Boomen DJH, Sienkiewicz A, Berlin I, Jongsma MLM, van Elsland DM, Luzio JP, et al. A trimeric Rab7 GEF controls NPC1-dependent lysosomal cholesterol export. Nat Commun. 2020; 11 (1):5559. Epub 20201103. https://doi.org/10.1038/s41467-020-19032-0 PMID: 33144569; PubMed Central PMCID: PMC7642327.

80. Patterson MC. A riddle wrapped in a mystery: understanding Niemann-Pick disease, type C. Neurologist. 2003; 9(6):301–10. https://doi.org/10.1097/01.nrl.0000094627.78754.5b PMID: 14629784.

81. Vanier MT. Niemann-Pick diseases. Handb Clin Neurol. 2013; 113:1717–21. https://doi.org/10.1016/B978-0-444-59565-2.00041-1 PMID: 23622394.

82. Bulik-Sullivan BK, Loh PR, Finucane HK, Ripke S, Yang J, Patterson N, et al. LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. Nat Genet. 2015; 47 (3):291–5. Epub 20150202. https://doi.org/10.1038/ng.3211 PMID: 25642630; PubMed Central PMCID: PMC4495769.

83. Cornelis MC, Tordoff MG, El-Sohemy A, van Dam RM. Recalled taste intensity, liking and habitual intake of commonly consumed foods. Appetite. 2017; 109:182–9. Epub 20161130. https://doi.org/10.1016/j.appet.2016.11.036 PMID: 27915079; PubMed Central PMCID: PMC5206898.

84. Ong J-S, Hwang L-D, Zhong VW, An J, Gharahkhani P, Breslin PAS, et al. Understanding the role of bitter taste perception in coffee, tea and alcohol consumption through Mendelian randomization. Scientific Reports. 2018; 8(1):16414. https://doi.org/10.1038/s41598-018-34713-z PMID: 30442986

85. Lipchock SV, Spielman AI, Mennella JA, Mansfield CJ, Hwang LD, Douglas JE, et al. Caffeine Bitterness is Related to Daily Caffeine Intake and Bitter Receptor mRNA Abundance in Human Taste Tissue. Perception. 2017; 46(3–4):245–56. Epub 20170124. https://doi.org/10.1177/0301006616686098 PMID: 28118781; PubMed Central PMCID: PMC5972367.

86. Masi C, Dinnella C, Pirastu N, Prescott J, Monteleone E. Caffeine metabolism rate influences coffee perception, preferences and intake. Food Quality and Preference. 2016; 53:97–104. https://doi.org/10.1016/j.foodqual.2016.06.002.

87. Cornelis MC, Byrne EM, Esko T, Nalls MA, Ganna A, Paynter N, et al. Genome-wide meta-analysis identifies six novel loci associated with habitual coffee consumption. Mol Psychiatry. 2015; 20(5):647–56. Epub 20141007. https://doi.org/10.1038/mp.2014.107 PMID: 25288136; PubMed Central PMCID: PMC4388784.

88. Cornelis MC, van Dam RM. Genetic determinants of liking and intake of coffee and other bitter foods and beverages. Scientific Reports. 2021; 11(1):23845. https://doi.org/10.1038/s41598-021-03153-7 PMID: 34903748

89. Edenberg HJ, Gelernter J, Agrawal A. Genetics of Alcoholism. Curr Psychiatry Rep. 2019; 21(4):26. Epub 20190309. https://doi.org/10.1007/s11920-019-1008-1 PMID: 30852706.

90. Kothapalli KS, Ye K, Gadgil MS, Carlson SE, O'Brien KO, Zhang JY, et al. Positive Selection on a Regulatory Insertion-Deletion Polymorphism in FADS2 Influences Apparent Endogenous Synthesis of Arachidonic Acid. Mol Biol Evol. 2016. Epub 2016/05/18. https://doi.org/10.1093/molbev/msw049 [pii] PMID: 27188529.

91. Vaites LP, Paulo JA, Huttlin EL, Harper JW. Systematic Analysis of Human Cells Lacking ATG8 Proteins Uncovers Roles for GABARAPs and the CCZ1/MON1 Regulator C18orf8/RMC1 in Macroautophagic and Selective Autophagic Flux. Mol Cell Biol. 2018; 38(1). Epub 2017/10/19. https://doi.org/10.1128/MCB.00392-17 PMID: 29038162; PubMed Central PMCID: PMC5730722.

92. Dean N, Zhang YB, Poster JB. The VRG4 gene is required for GDP-mannose transport into the lumen of the Golgi in the yeast, Saccharomyces cerevisiae. J Biol Chem. 1997; 272(50):31908–14. https://doi.org/10.1074/jbc.272.50.31908 PMID: 9395539.

93. Rodríguez A, Gonzalez L, Ko A, Alvarez M, Miao Z, Bhagat Y, et al. Molecular Characterization of the Lipid Genome-Wide Association Study Signal on Chromosome 18q11.2 Implicates HNF4A-Mediated Regulation of the TMEM241 Gene. Arterioscler Thromb Vasc Biol. 2016; 36(7):1350–5. Epub 2016/05/21. https://doi.org/10.1161/ATVBAHA.116.307182 PMID: 27199446; PubMed Central PMCID: PMC5154300.

94. Sugawara T. Sphingolipids as Functional Food Components: Benefits in Skin Improvement and Disease Prevention. J Agric Food Chem. 2022; 70(31):9597–609. Epub 20220729. https://doi.org/10.1021/acs.jafc.2c01731 PMID: 35905137.

95. Wang X, Wang Y, Xu J, Xue C. Sphingolipids in food and their critical roles in human health. Crit Rev Food Sci Nutr. 2021; 61(3):462–91. Epub 20200325. https://doi.org/10.1080/10408398.2020.1736510 PMID: 32208869.

96. Okuda T. Dietary Control of Ganglioside Expression in Mammalian Tissues. Int J Mol Sci. 2019; 21(1). Epub 20191226. https://doi.org/10.3390/ijms21010177 PMID: 31887977; PubMed Central PMCID: PMC6981639.

97. Li W, Belwal T, Li L, Xu Y, Liu J, Zou L, et al. Sphingolipids in foodstuff: Compositions, distribution, digestion, metabolism and health effects—A comprehensive review. Food Res Int. 2021; 147:110566. Epub 20210701. https://doi.org/10.1016/j.foodres.2021.110566 PMID: 34399542.

98. Colombaioni L, Garcia-Gil M. Sphingolipid metabolites in neural signalling and function. Brain Res Brain Res Rev. 2004; 46(3):328–55. https://doi.org/10.1016/j.brainresrev.2004.07.014 PMID: 15571774.

99. Hussain G, Wang J, Rasul A, Anwar H, Imran A, Qasim M, et al. Role of cholesterol and sphingolipids in brain development and neurological diseases. Lipids Health Dis. 2019; 18(1):26. Epub 20190125. https://doi.org/10.1186/s12944-019-0965-z PMID: 30683111; PubMed Central PMCID: PMC6347843.

100. Kulminski AM, Loiko E, Loika Y, Culminskaya I. Pleiotropic predisposition to Alzheimer's disease and educational attainment: insights from the summary statistics analysis. Geroscience. 2022; 44(1):265–80. Epub 20211106. https://doi.org/10.1007/s11357-021-00484-1 PMID: 34743297; PubMed Central PMCID: PMC8572080.

101. Genomic Relationships, Novel Loci, and Pleiotropic Mechanisms across Eight Psychiatric Disorders. Cell. 2019; 179(7):1469–82.e11. https://doi.org/10.1016/j.cell.2019.11.020 PMID: 31835028; PubMed Central PMCID: PMC7077032.

102. Perry CL, McGuire MT, Neumark-Sztainer D, Story M. Characteristics of vegetarian adolescents in a multiethnic urban population. J Adolesc Health. 2001; 29(6):406–16. https://doi.org/10.1016/s1054-139x(01)00258-0 PMID: 11728890.

103. McLean CP, Moeck EK, Sharp G, Thomas NA. Characteristics and clinical implications of the relationship between veganism and pathological eating behaviours. Eat Weight Disord. 2022; 27(5):1881–6. Epub 20211116. https://doi.org/10.1007/s40519-021-01330-1 PMID: 34786670.

104. Klopp SA, Heiss CJ, Smith HS. Self-reported vegetarianism may be a marker for college women at risk for disordered eating. J Am Diet Assoc. 2003; 103(6):745–7. https://doi.org/10.1053/jada.2003.50139 PMID: 12778048.

105. Sieke EH, Carlson JL, Lock J, Timko CA, Neumark-Sztainer D, Peebles R. To meat or not to meat: disordered eating and vegetarian status in university students. Eat Weight Disord. 2022; 27(2):831–7. Epub 20210522. https://doi.org/10.1007/s40519-021-01202-8 PMID: 34021903.

106. Hansson LM, Björck C, Birgegård A, Clinton D. How do eating disorder patients eat after treatment? Dietary habits and eating behaviour three years after entering treatment. Eat Weight Disord. 2011; 16(1):e1–8. https://doi.org/10.1007/BF03327514 PMID: 21727776.

107. Yackobovitch-Gavan M, Golan M, Valevski A, Kreitler S, Bachar E, Lieblich A, et al. An integrative quantitative model of factors influencing the course of anorexia nervosa over time. Int J Eat Disord. 2009; 42(4):306–17. https://doi.org/10.1002/eat.20624 PMID: 19040269.

108. Bakan R, Birmingham CL, Aeberhardt L, Goldner EM. Dietary zinc intake of vegetarian and nonvegetarian patients with anorexia nervosa. Int J Eat Disord. 1993; 13(2):229–33. https://doi.org/10.1002/1098-108x(199303)13:2<229::aid-eat2260130211>3.0.co;2-1 PMID: 8477292.

109. Fry A, Littlejohns TJ, Sudlow C, Doherty N, Adamska L, Sprosen T, et al. Comparison of Sociodemographic and Health-Related Characteristics of UK Biobank Participants With Those of the General Population. Am J Epidemiol. 2017; 186(9):1026–34. https://doi.org/10.1093/aje/kwx246 PMID: 28641372; PubMed Central PMCID: PMC5860371.