Pro-inflammatory fatty acid profile and colorectal cancer risk

May-Wilson, Sebastian

2017-10


http://hdl.handle.net/10138/225155
https://doi.org/10.1016/j.ejca.2017.07.034

Downloaded from Helda, University of Helsinki institutional repository.

This is an electronic reprint of the original article.

This reprint may differ from the original in pagination and typographic detail.

Please cite the original version.
Pro-inflammatory fatty acid profile and colorectal cancer risk: A Mendelian randomisation analysis

Sebastian May-Wilson a,1, Amit Sud a,1, Philip J. Law a,1, Kimmo Palin b,c, Sari Tuupanen b,c, Alexandra Gylfe b,c, Ulrika A. Hänninen b,c, Tatiana Cajuso b,c, Tomas Tanskanen b,c, Johanna Kondelin b,c, Eevi Kaasinen b,c, Anti-Pekka Sarin d, Johan G. Eriksson g,h,i, Harri Rissanen g, Paul Knek t g, Eero Pukkala j,k, Pekka Jousilahti g, Veikko Salomaa g, Samuli Ripatti d,e,f, Aarno Palotie d,l,m,n, Laura Renkonen-Sinisalo o, Anna Lepistö o, Jan Böhm p, Jukka-Pekka Mecklin q, Nada A. Al-Tassan r, Claire Palles s, Susan M. Farrington t, Maria N. Timofeeva t, Brian F. Meyer r, Salma M. Waki l, Harry Campbell u, Christopher G. Smith v, Shelley Idziaszczyk v, Timothy S. Maughan w, David Fisher x, Rachel Kerr y, David Kerr z, Michael N. Passarelli aa, Jane C. Figueiredo ab,ac, Daniel D. Buchanan ad,ae, Aung K. Win ae, John L. Hopper ae, Mark A. Jenkins ae, Noralane M. Lindor af, Polly A. Newcomb ag, Steven Gallinger ah, David Conti ai, Fred Schumacher aj, Graham Casey aj, Lauri A. Aaltonen b,c, Jeremy P. Cheadle v, Ian P. Tomlinson s, Malcolm G. Dunlop t, Richard S. Houlston a,*

a Division of Genetics and Epidemiology, The Institute of Cancer Research, London, SW7 3RP, UK
b Genome-Scale Biology Research Program, Research Programs Unit, University of Helsinki, Helsinki, 00014, Finland
c Department of Medical and Clinical Genetics, Medicum, University of Helsinki, Helsinki, 00014, Finland
d Institute for Molecular Medicine Finland (FIMM), University of Helsinki, Helsinki, 00014, Finland
e Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, CB10 1SA, UK
f Department of Public Health, University of Helsinki, Helsinki, 00014, Finland
gh Unit of General Practice and Primary Health Care, University of Helsinki and Helsinki University Hospital, Helsinki, 00014, Finland
i National Institute for Health and Welfare, Helsinki, 00271, Finland
j Folkhälsan Research Centre, Helsinki, 00250, Finland

* Corresponding author: Division of Genetics and Epidemiology, The Institute of Cancer Research, 15 Cotswold Road, London, SM2 5NG, UK.
Fax: +44 (0) 722 4365.
E-mail address: richard.houlston@icr.ac.uk (R.S. Houlston).

Available online at www.sciencedirect.com

http://dx.doi.org/10.1016/j.ejca.2017.07.034
0959-8049/© 2017 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).
**KEYWORDS**
Mendelian randomisation; Colorectal cancer; Risk; Plasma fatty acids; Fatty acids

**Abstract**

**Background:** While dietary fat has been established as a risk factor for colorectal cancer (CRC), associations between fatty acids (FAs) and CRC have been inconsistent. Using Mendelian randomisation (MR), we sought to evaluate associations between polyunsaturated (PUFA), monounsaturated (MUFA) and saturated FAs (SFAs) and CRC risk.

**Methods:** We analysed genotype data on 9254 CRC cases and 18,386 controls of European ancestry. Externally weighted polygenic risk scores were generated and used to evaluate associations with CRC per one standard deviation increase in genetically defined plasma FA levels.

**Results:** Risk reduction was observed for oleic and palmitoleic MUFAs (OR OA = 0.77, 95% CI: 0.65–0.92, P = 3.9 × 10⁻³; OR POA = 0.36, 95% CI: 0.15–0.84, P = 0.018). PUFA linoleic and arachidonic acid had negative and positive associations with CRC respectively (OR LA = 0.95, 95% CI: 0.93–0.98, P = 3.7 × 10⁻⁴; OR AA = 1.05, 95% CI: 1.02–1.07, P = 1.7 × 10⁻⁴). The SFA stearic acid was associated with increased CRC risk (OR SA = 1.17, 95% CI: 1.01–1.35, P = 0.041).

**Conclusion:** Results from our analysis are broadly consistent with a pro-inflammatory FA profile having a detrimental effect in terms of CRC risk.

© 2017 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).
1. Introduction

Colorectal cancer (CRC) is one of the most common cancers and a major cause of cancer-related mortality in economically developed countries [1]. Geographical differences in CRC incidence between countries and migration studies have established the importance of lifestyle and diet as major determinants for CRC risk [2]. Worldwide CRC is currently diagnosed in over one million individuals annually; however, its incidence is set to increase with adoption of western lifestyles in developing countries [3]. Given the importance of diet as a risk factor for CRC, its modification offers the prospect of impacting significantly on disease incidence through public health initiatives.

Dietary fat has been widely implicated as a risk factor for cancer, and meta-analyses of epidemiological studies have tended to associate CRC risk with a higher consumption of red and processed meat [4]. The association between fat intake on cancer risk however, is likely to depend not only on the quantity, but also on the specific type of fatty acid (FA). Animal models and ecological studies have tended to implicate animal fat [5], saturated fatty acid (SFA) and certain omega-6 polyunsaturated fatty acids (ω-6 PUFAs) with an increased risk, and ω-3 PUFA intake with a reduced risk [6–8]. Evidence for a causal relationship with intake of specific types of fat from epidemiological studies has however largely been inconclusive. Reasons for inconsistencies in observational studies include the inherent problem of eliciting accurate measurements of long-term diet, confounding and reverse causation [9].

Mendelian randomisation (MR) analysis represents an adjunct to the conventional epidemiological observational study for examining associations between an exposure with a disease. The MR strategy makes use of allelic variants that are randomly assigned during meiosis and are robustly associated with traits of interest, as instrumental variables (IVs). Using genetically defined IVs as proxies of modifiable exposure avoids reverse causality and can inform on life-long exposure with a disease. The MR strategy makes use of genetic risk scores (Table 1). Briefly, we excluded single nucleotide polymorphisms (SNPs) with a minor allele frequency of <1%, low call rate <95%, those SNPs violating Hardy–Weinberg equilibrium, and individuals with non-European ancestry as assessed using data from HapMap v2 [15]. IMPUTEv2 software [16] was used to recover untyped SNP genotypes using a merged reference panel consisting of Sequencing Initiative Suomi (for the FINLAND data) or UK10K (for the remaining data) and 1000 Genomes Project data [17,18]. Poorly imputed SNPs, defined by an INFO score of <0.9, were excluded. Summary statistics from the seven GWAS were used to calculate the odds ratios (ORs) for FA-related SNPs.

2. Methods

2.1. Colorectal cancer datasets

We investigated the relationship between genetic risk scores for levels of MUFAs, PUFAs, and SFAs and CRC risk adopting a two-sample MR strategy using data from seven reported genome-wide association studies (GWAS) of CRC (Table 1). Briefly, these GWAS were based on individuals with European ancestry: CCFR1, CCFR2, COIN, FINLAND, UK1, Scotland1 and VQ58 [14]. Each study was approved by respective institutional ethics review board and performed/conducted in accordance with the Declaration of Helsinki.

2.2. Genotyping data

Comprehensive details of the genotyping and quality control of the seven GWAS have been previously reported [14]. Briefly, we excluded single nucleotide polymorphisms (SNPs) with a minor allele frequency of <1%, low call rate <95%, those SNPs violating Hardy–Weinberg equilibrium, and individuals with non-European ancestry as assessed using data from HapMap v2 [15]. IMPUTEv2 software [16] was used to recover untyped SNP genotypes using a merged reference panel consisting of Sequencing Initiative Suomi (for the FINLAND data) or UK10K (for the remaining data) and 1000 Genomes Project data [17,18]. Poorly imputed SNPs, defined by an INFO score of <0.9, were excluded. Summary statistics from the seven GWAS were used to calculate the odds ratios (ORs) for FA-related SNPs.

2.3. Gene variants used to construct genetic risk scores

Genetic risk scores for IVs for each plasma FA were developed from SNPs previously identified by The Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium. We considered SNPs associated at genome-wide significance (i.e. $P \leq 5.0 \times 10^{-8}$) in individuals with European Ancestry. To avoid co-linearity between SNPs for each FA we imposed a threshold $r^2$ value of ≥0.01 for linkage disequilibrium (LD) including only the SNPs with the strongest effect on the trait in genetic risk scores (Table 2, [19–22]). For each identified SNP, we recovered the chromosome positions, the risk alleles, association estimates and standard errors. For each SNP, the allele that was associated with increased FA level was considered the effect allele.

2.4. Statistical analysis

The association between the plasma level of each FA and CRC was examined using MR on summary
Table 1
Summary of the seven colorectal cancer genome-wide association studies.

<table>
<thead>
<tr>
<th>Series</th>
<th>Study setting</th>
<th>Study centre</th>
<th>Genotyping platform</th>
<th>No. cases</th>
<th>No. controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>CCFR1</td>
<td>Colon Cancer Family Registry</td>
<td>University of Southern California</td>
<td>Illumina 1M, 1M Duo</td>
<td>1290</td>
<td>1055</td>
</tr>
<tr>
<td>CCFR2</td>
<td>Colon Cancer Family Registry</td>
<td>University of Southern California</td>
<td>Illumina 1M, Omni express</td>
<td>796</td>
<td>2236</td>
</tr>
<tr>
<td>COIN</td>
<td>COIN trial: Multicentre study of cetuximab and other therapies in metastatic CRC. Controls were unselected blood donors</td>
<td>Cardiff University</td>
<td>Affymetrix Axiom</td>
<td>2244</td>
<td>2162</td>
</tr>
<tr>
<td>FINLAND</td>
<td>Finnish Colorectal Cancer Predisposition Study</td>
<td>Helsinki University</td>
<td>Illumina 610K/Illumina HumanOmni2.5M</td>
<td>1172</td>
<td>8266</td>
</tr>
<tr>
<td>UK1</td>
<td>CORGI (colorectal Tumour Gene Identification Consortium)</td>
<td>Oxford University</td>
<td>Illumina Hap550</td>
<td>940</td>
<td>965</td>
</tr>
<tr>
<td>Scotland1</td>
<td>COGS (Colorectal Cancer Susceptibility Study)</td>
<td>Edinburgh University</td>
<td>Illumina Hap300/240S</td>
<td>1012</td>
<td>1012</td>
</tr>
<tr>
<td>VQ58</td>
<td>Cases: VICTOR, post-treatment stages of a phase III, randomised trial of rofecoxib (VIOXX) in patients after potentially curative therapy. QUASAR2, multi-centre study of capecitabine + bevacizumab as adjuvant treatment. 1958 Birth cohort controls</td>
<td>Oxford University</td>
<td>Illumina Hap300/370, Illumina 1M</td>
<td>1800</td>
<td>2690</td>
</tr>
</tbody>
</table>

Table 2
Effect sizes for plasma fatty acid content (per standard deviation increase in levels) for genome-wide significant \( (P < 5 \times 10^{-8}) \) instrumental variables reported by CHARGE consortium.

<table>
<thead>
<tr>
<th>FA subtype</th>
<th>Fatty acid</th>
<th>SNP ID</th>
<th>Chr</th>
<th>Position (bp)</th>
<th>Allele</th>
<th>( \beta )</th>
<th>StdErr</th>
<th>( P )-value</th>
<th>Variance explained*</th>
</tr>
</thead>
<tbody>
<tr>
<td>SFA</td>
<td>Arachidic acid (20:0)</td>
<td>rs680379</td>
<td>20</td>
<td>12917400</td>
<td>A/G</td>
<td>0.098</td>
<td>0.01</td>
<td>5.81 \times 10^{-13}</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>Palmitic acid (PA) (16:0)</td>
<td>rs2391388</td>
<td>1</td>
<td>95485825</td>
<td>C/A</td>
<td>0.18</td>
<td>0.03</td>
<td>2.72 \times 10^{-11}</td>
<td>0.21–0.98%</td>
</tr>
<tr>
<td></td>
<td>Stearic acid (SA) (18:0)</td>
<td>rs6675668</td>
<td>1</td>
<td>95515637</td>
<td>G/T</td>
<td>0.17</td>
<td>0.02</td>
<td>2.16 \times 10^{-18}</td>
<td>0.37–1.39%</td>
</tr>
<tr>
<td></td>
<td>Docosahexaenoic acid (DHA) (22:6n-3)</td>
<td>rs11119805</td>
<td>1</td>
<td>211918244</td>
<td>T/A</td>
<td>0.17</td>
<td>0.03</td>
<td>2.8 \times 10^{-8}</td>
<td>&lt;0.01–0.72</td>
</tr>
<tr>
<td></td>
<td>Docosapentaenoic acid (DPA) (22:5n-3)</td>
<td>rs102275</td>
<td>11</td>
<td>61557803</td>
<td>T/C</td>
<td>0.18</td>
<td>0.02</td>
<td>1.33 \times 10^{-20}</td>
<td>0.33–1.34%</td>
</tr>
<tr>
<td>( \omega-3 ) PUFA</td>
<td>Eicosapentaenoic acid (EPA) (20:5n-3)</td>
<td>rs22362126</td>
<td>6</td>
<td>10995015</td>
<td>G/C</td>
<td>0.11</td>
<td>0.01</td>
<td>1.26 \times 10^{-15}</td>
<td>0.7%</td>
</tr>
<tr>
<td>( \omega-6 ) PUFA</td>
<td>Arachidonic acid (AA) (20:4n-6)</td>
<td>rs174547</td>
<td>11</td>
<td>61570783</td>
<td>T/C</td>
<td>1.69</td>
<td>0.03</td>
<td>3.3 \times 10^{-97}</td>
<td>3.7–37.6%</td>
</tr>
<tr>
<td></td>
<td>Dihomo-( \gamma )-linolenic acid (DGLA) (20:3n-6)</td>
<td>rs174547</td>
<td>11</td>
<td>61570783</td>
<td>C/T</td>
<td>0.36</td>
<td>0.01</td>
<td>2.63 \times 10^{-15}</td>
<td>8.7–11.1%</td>
</tr>
<tr>
<td></td>
<td>Linoleic acid (LA) (18:2n-6)</td>
<td>rs174547</td>
<td>11</td>
<td>61570783</td>
<td>C/T</td>
<td>1.47</td>
<td>0.04</td>
<td>4.98 \times 10^{-24}</td>
<td>7.6–18.1%</td>
</tr>
<tr>
<td>( \omega-7 ) MUFA</td>
<td>Palmitoleic acid (POA) (16:1n-7)</td>
<td>rs780093</td>
<td>2</td>
<td>27742603</td>
<td>T/C</td>
<td>0.02</td>
<td>0.003</td>
<td>9.80 \times 10^{-10}</td>
<td>0.23–0.93%</td>
</tr>
<tr>
<td></td>
<td>Oleic acid (OA) (18:1n-9)</td>
<td>rs6722456</td>
<td>2</td>
<td>134529091</td>
<td>G/A</td>
<td>0.05</td>
<td>0.009</td>
<td>4.12 \times 10^{-8}</td>
<td>&lt;0.01–0.57</td>
</tr>
</tbody>
</table>

FA, fatty acid; SNP, single nucleotide polymorphism; bp, base pair; SFA, saturated fatty acid; MUFA, monounsaturated fatty acid; PUFA, polyunsaturated fatty acid; StdErr, standard error. Effect allele influencing each FA trait is marked in bold.

*a* hg19 NCBI build.

* Taken from CHARGE consortium, as a percentage of total serum fatty acids, calculated by \((\beta^2 \times \text{MAF}^2 \times (1-\text{MAF})/\text{Var}(Y))\) where \( \beta \) is the regression coefficient, MAF is the minor allele frequency and \( \text{Var}(Y) \) is the variance in levels of the fatty acid. IVs obtained from Refs. [19,20,22].
statistics as per Burgess (2015) [23]. The ratio estimate ($\hat{\beta}$) of all SNPs associated with each fatty acid, combined, on CRC was calculated as follows:

$$\hat{\beta} = \frac{\sum X_k^2 Y_k \sigma_{Y_k}^2}{\sum X_k^2 \sigma_{Y_k}^2}.$$ 

where $X_k$ corresponds to the association of SNP $k$ (as log of the OR per risk allele) with the fatty acid trait $Y_k$; $Y_k$ is the association between SNP $k$ and CRC risk (as log of the OR) with standard error $\sigma_{Y_k}$. The estimate for ($\hat{\beta}$) represents the causal increase in the log odds of the CRC, per unit change in fatty acids. The standard error of the combined ratio estimate is given by:

$$se(\hat{\beta}) = \sqrt{\frac{1}{\sum X_k^2 \sigma_{Y_k}^2}}.$$ 

A meta-analysis of statistics for each specific FA generated for each CRC cohort was combined under fixed-effects models to derive the summary ORs and confidence intervals (CIs). To assess the impact of between-study heterogeneity, we also derived ORs under a random-effects model.

A central tenet in MR is the absence of pleiotropy (i.e. a gene influencing multiple traits) between the SNPs influencing CRC risk and FA levels. This would be revealed as deviation from a linear relationship between SNPs and their effect size for any FA and CRC risk. To examine for violation of the standard IV assumptions in our analysis, we performed inverse variant weighted (IVW) and MR-Egger regression tests [24].

We considered a significance level of $P \leq 0.05$ as being satisfactory to derive a conclusion. While ordinarily it would be appropriate to impose a Bonferroni-corrected threshold, this assumes an independence of IVs across all FA traits, which is not the case in the present analysis. All statistical analyses were undertaken using R version 3.1 software [25].

### 2.5. Expression quantitative trait locus analysis

To examine the relationship between SNP genotype and expression of FA metabolism genes, we performed expression quantitative trait locus (eQTL) analysis using data from The Cancer Genome Atlas (TCGA) and the genotype tissue expression (GTEx) project [26,27].

### 3. Results

The FA-associated genetic variants and their GWAS-reported characteristics that were used to derive IVs for FAs are detailed in Table 2. A reduced risk of CRC was observed for genetic variants associated with increases in the MUFAs studied (Table 3). In all but one of the seven cohorts increased levels of OA were associated with reduced CRC risk (Fig. 1). In the meta-analysis of these seven cohorts the OR$_{OA}$ was 0.77 (95% CI: 0.65–0.92, $P = 3.9 \times 10^{-3}$) with little evidence of between-study heterogeneity ($P_{het} = 0.23, I^2 = 26\%$). Similarly, increased levels of POA were associated with reduced CRC risk with an OR$_{POA}$ of 0.36 (95% CI: 0.15–0.84, $P = 0.018, P_{het} = 0.08, I^2 = 47\%$; Fig. 1).

The $\omega$-6 PUFAs LA and AA both showed association with CRC risk, but in different directions. Specifically, LA was associated with reduced risk (OR$_{LA} = 0.95, 95\%$ CI: 0.93–0.98, $P = 3.7 \times 10^{-4}$, $P_{het} = 0.03, I^2 = 57\%$; Fig. 1) and AA with an increased risk (OR$_{AA} = 1.05, 95\%$ CI: 1.02–1.07, $P = 1.7 \times 10^{-4}$, $P_{het} = 0.03, I^2 = 56\%$). The association between one standard deviation increase in each of the other PUFAs defined by their respective IVs and CRC risk were null (Supplementary Fig. 1).

Of the three SFAs studied, increased SA was nominally associated with CRC risk (OR$_{SA} = 1.17, 95\%$ CI: 1.01–1.35, $P = 0.041, P_{het} = 0.04, I^2 = 55\%$). To formally assess the impact of heterogeneity on study findings we derived ORs under a random-effects

<table>
<thead>
<tr>
<th>Fatty acid</th>
<th>Significant associations</th>
<th>OR (fixed effects)</th>
<th>95% CI (fixed effects)</th>
<th>$P$-value (fixed effects)</th>
<th>OR (random effects)</th>
<th>95% CI (random effects)</th>
<th>$P$-value (random effects)</th>
<th>$I^2$</th>
<th>$P_{het}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arachidonic acid</td>
<td></td>
<td>0.92</td>
<td>0.61–1.39</td>
<td>0.7</td>
<td>0.93</td>
<td>0.61–1.40</td>
<td>0.71</td>
<td>3%</td>
<td>0.41</td>
</tr>
<tr>
<td>Palmitic acid (PA)</td>
<td></td>
<td>0.97</td>
<td>0.78–1.21</td>
<td>0.82</td>
<td>0.97</td>
<td>0.78–1.21</td>
<td>0.82</td>
<td>0%</td>
<td>0.47</td>
</tr>
<tr>
<td>Serronic acid (SA)</td>
<td></td>
<td>1.16</td>
<td>1.01–1.35</td>
<td>0.04</td>
<td>1.2</td>
<td>0.95–1.49</td>
<td>0.12</td>
<td>55%</td>
<td>0.04</td>
</tr>
<tr>
<td>Docosahexaenoic acid (DHA)</td>
<td></td>
<td>1.32</td>
<td>0.94–1.87</td>
<td>0.11</td>
<td>1.32</td>
<td>0.94–1.87</td>
<td>0.11</td>
<td>0%</td>
<td>0.65</td>
</tr>
<tr>
<td>Docosapentaenoic acid (DPA)</td>
<td></td>
<td>1.58</td>
<td>0.99–2.52</td>
<td>0.06</td>
<td>1.63</td>
<td>0.97–2.73</td>
<td>0.06</td>
<td>17%</td>
<td>0.3</td>
</tr>
<tr>
<td>Eicosapentaenoic acid (EPA)</td>
<td></td>
<td>0.39</td>
<td>0.13–1.21</td>
<td>0.1</td>
<td>0.39</td>
<td>0.13–1.21</td>
<td>0.1</td>
<td>0%</td>
<td>0.57</td>
</tr>
<tr>
<td>Arachidonic acid (AA)</td>
<td></td>
<td>1.05</td>
<td>1.02–1.07</td>
<td>1.7 $\times 10^{-4}$</td>
<td>1.05</td>
<td>1.02–1.09</td>
<td>4.9 $\times 10^{-3}$</td>
<td>56%</td>
<td>0.03</td>
</tr>
<tr>
<td>Dihomo-γ-linolenic acid (GLA)</td>
<td></td>
<td>0.91</td>
<td>0.83–1.00</td>
<td>0.06</td>
<td>0.95</td>
<td>0.80–1.01</td>
<td>0.07</td>
<td>23%</td>
<td>0.26</td>
</tr>
<tr>
<td>Linoleic acid (LA)</td>
<td></td>
<td>0.95</td>
<td>0.93–0.98</td>
<td>3.7 $\times 10^{-4}$</td>
<td>0.95</td>
<td>0.91–0.99</td>
<td>8.9 $\times 10^{-3}$</td>
<td>57%</td>
<td>0.03</td>
</tr>
<tr>
<td>Oleic acid (OA)</td>
<td></td>
<td>0.77</td>
<td>0.65–0.92</td>
<td>3.9 $\times 10^{-3}$</td>
<td>0.76</td>
<td>0.62–0.94</td>
<td>9.7 $\times 10^{-3}$</td>
<td>26%</td>
<td>0.23</td>
</tr>
<tr>
<td>Palmitoleic acid (POA)</td>
<td></td>
<td>0.36</td>
<td>0.15–0.84</td>
<td>0.018</td>
<td>0.32</td>
<td>0.10–1.07</td>
<td>0.06</td>
<td>47%</td>
<td>0.08</td>
</tr>
</tbody>
</table>

$P_{het}$, $P$-value for heterogeneity; $I^2$, proportion of the total variation due to heterogeneity; SFA, saturated fatty acid; PUFA, polyunsaturated fatty acid; MUFA, monounsaturated fatty acid.
model. Associations between AA, LA and OA and CRC risk remained significant (Table 3).

We assessed the impact of possible classical pleiotropism on MR estimates using both IVW and MR-Egger regression tests. There was no evidence for violation of the standard IV assumptions used for MR analysis, such as a dependence on confounders (Table 4).

In the present analysis, we used the SNP rs102275 in combination with other SNPs to generate a polygenic risk score for SA, OA and POA, whereas rs174547, which is in LD with rs102275 \((r^2 = 1.0 \text{ and } D' = 1.0)\), was used for DPA, AA, DGLA and LA. Both SNPs annotate the \(FADS2\) gene. \(FADS2\) is a rate-limiting enzyme in the desaturation of LA to AA, and \(\alpha\)-linolenic acid into DHA and EPA (Fig. 2). These FAs are precursors for prostaglandins and leukotrienes, which are key mediators of the inflammatory response. In an eQTL analysis rs174547 and rs102275 genotype were shown to be strongly correlated with \(FADS2\) expression across a range of different tissue types, including blood \((P = 3.98 \times 10^{-10})\), normal colon \((P = 1.65 \times 10^{-10})\) and CRC \((P = 2.07 \times 10^{-5})\) (Supplementary Table 1).

---

**Fig. 1.** Meta-analysis odds ratios (OR) for colorectal cancer per unit increase in genetic risk score (standard deviation of trait) for significant fatty acid associations. (a) Oleic acid; (b) arachidonic acid; (c) stearic acid; (d) linoleic acid; (e) palmitoleic acid; \(I^2\): proportion of the total variation due to heterogeneity. Boxes: OR point estimate; its area is proportional to the weight of the study. Diamond: overall summary estimate, with confidence intervals given by its width. Vertical line: null value (OR = 1.0).
Table 4
IVW and MR-Egger test results for combined fatty acid instrumental variables.

<table>
<thead>
<tr>
<th>Fatty acid subtype</th>
<th>Fatty acid</th>
<th>IVW</th>
<th>MR-Egger</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Slope Estimate</td>
<td>P-value</td>
</tr>
<tr>
<td>SFA</td>
<td>Stearic acid (SA)</td>
<td>$-0.1$ (−0.33 to 0.64)</td>
<td>0.30</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Slope</td>
<td></td>
</tr>
<tr>
<td>ω-3 PUFA</td>
<td>Docosapentaenoic acid (DPA)</td>
<td>$0.46$ (−2.32 to 3.23)</td>
<td>0.55</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Slope</td>
<td></td>
</tr>
<tr>
<td>ω-6 PUFA</td>
<td>Arachidonic acid (AA)</td>
<td>$-0.59$ (−7.99 to 9.16)</td>
<td>0.54</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Slope</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Linoleic acid (LA)</td>
<td>$-0.09$ (−2.48 to 2.29)</td>
<td>0.70</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Slope</td>
<td></td>
</tr>
<tr>
<td></td>
<td>α-linolenic acid (DGLA)</td>
<td>$-0.05$ (−0.17 to 0.07)</td>
<td>0.22</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Slope</td>
<td></td>
</tr>
<tr>
<td>MUFA</td>
<td>Palmitoleic acid (POA)</td>
<td>$-1.03$ (−2.64 to 0.58)</td>
<td>0.15</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Slope</td>
<td></td>
</tr>
</tbody>
</table>

CI, confidence interval; MUFA, monounsaturated fatty acid; PUFA, polyunsaturated fatty acid; SFA, saturated fatty acid; IVW, inverse variant weighted. *FA traits with two IVs, preventing calculation of CIs and P-value.

4. Discussion

While dietary fat intake has been associated with the CRC risk, teasing out specific FA associations and their mechanistic basis has proven to be challenging. A number of observational studies have reported associations between serum levels of specific FAs with CRC [28,29], supporting our findings.

A major strength of the MR strategy to identify causal associations is that it is not influenced by recall bias and confounding that can affect traditional observational studies. Nevertheless, a key assumption in MR is that the variants used to generate genetic scores are associated with the exposure being queried. Herein, we only made use of SNPs associated with each FA at genome-wide significance from hypothesis-free GWAS. Furthermore, we only used data from individuals of European descent so as to limit bias from population stratification. Another central assumption in MR is that variants are associated with CRC only through the exposure and are not confounded by pleiotropy, which would be revealed by a positive correlation between increasing effect sizes in the IVs and CRC risk. While we did not observe such relationship, we acknowledge that IVs for a number of the FAs were solely based on only one or two SNPs, preventing assessment by IVW and MR-Egger analysis. One strategy to overcome this and fully investigate any pleiotropy would be to measure FA serum levels in correlation with CRC risk.

In this analysis, the same SNP (rs102275, or correlated SNP rs174547) was used to make causal deductions between multiple FAs and CRC risk. Therefore, SNPs have been used each time assuming that the exposure individually accounts for the disease association. The genetic variant association with CRC risk is consequently double-counted, in that the effect is attributed to different FA exposures [30]. With such vertical pleiotropism, single locus MR analyses cannot robustly decipher which FA is primarily driving the relationship with CRC risk. Such considerations have not been addressed in previous studies of the relationship between PUFAs and prostate cancer [31] or between branched-chain amino acids and diabetes [32].

While we did not demonstrate a causal association between other FAs including several PUFAs, SFAs and CRC risk, we acknowledge that our power to demonstrate a relationship was limited. For example, with respect to EPA: assuming the variance explained by the alleles is 0.04%, based on epidemiological observational study data, and a relative risk of 1.04 we had <10% power to demonstrate a relationship [33].

Accepting these caveats we have provided support for differing effects of OA, and ω-6 PUFAs LA and AA on CRC risk. Our findings broadly accord with the findings from many of the published ecological and epidemiological observational studies. Notably, increased levels of AA contribute as a risk factor to CRC development [34,35], while increased intake of olive oil, which is high in OA, is associated with decreased risk [36]. A number of epidemiological studies have provided evidence that a Mediterranean diet, with a higher olive oil intake, is associated with reduced CRC risk [36–38].

In the eQTL analysis, both rs102275 and rs174547 show evidence of cis-regulatory effects on FADS2 expression. Intriguingly, rs174547 has previously been reported to have opposing effects on FADS2 and FADSI expression in CRC [39]. Collectively, these data provide for relationship between diet, genotype, FA metabolism and CRC risk through modulation of an inflammatory response.

Even so, a biological basis for associations between specific FAs and CRC risk remain to be established. It is however, predicted a priori that within any FA class, different members have different actions and effects. With
respects to ω-6, evidence supports the inflammatory effects for AA through COX-2 production of inflammatory mediators [40] including prostaglandin E2, which affect CRC carcinogenesis [41–43]. This implies that diets high in AA, such as meat or eggs, may lead to more inflammatory compounds, which in turn may increase CRC risk. While increasing dietary LA, an essential FA, might potentially enrich tissues with AA due to their metabolic link [44], a gene–environment interaction may exist to influence colon FA content [45]. There is however, contradictory evidence from studies that have associated LA with both an increased [46] and decreased risk of CRC, possibly by altering ω-6 to ω-3 FA ratios [47] or alternatively production of reactive oxygen species [48]. The ability of aspirin to irreversibly inhibit COX-1 and COX-2 and therefore lower pro-inflammatory signals independent of genotype and diet, has thus proved an attractive option for CRC chemoprevention [49].

Fig. 2. Pathway of fatty acids. Shown are the various fatty acids analysed, and the enzymes involved in their metabolism. COX: cyclooxygenase, LOX: 5-lipoxygenase.
In conclusion, irrespective of the biological basis of associations between FAs and CRC risk our findings are consistent with the observation that the dietary composition of MUFAs in Mediterranean diets are risk reducing, and that a pro-inflammatory diet are risk increasing [50]. While we may not be at a stage where we can justifiably advise individuals to alter their intake of specific FAs to decrease the risk of developing CRC, it seems the current guidelines to moderate total fat and SFA consumption and increase unsaturated FA intake is likely to be beneficial.

**Conflict of interest statement**

None declared.

**Acknowledgements**

At the Institute of Cancer Research, this work was supported by Cancer Research UK (C1298/A8362 - Bobby Moore Fund for Cancer Research UK). Additional support was provided by the National Cancer Research Network. S. M-W was in receipt of a PhD studentship from The Institute of Cancer Research. A.S. is supported by a clinical fellowship from Cancer Research UK. In Edinburgh the work was supported by Programme Grant funding from Cancer Research UK (C348/A12076). In Oxford additional funding was provided by the Oxford Comprehensive Biomedical Research Centre and the EU FP7 CHIBCHA grant. Core infrastructure support to the Wellcome Trust Centre for Human Genetics, Oxford was provided by grant (090532/Z/09/Z). We are grateful to many colleagues within UK Clinical Genetics Departments (for CORGI) and to many collaborators who participated in the VICTOR and QUASAR2 trials. We also thank colleagues from the UK National Cancer Research Network (for NSCCG). Support from the European Union (FP7/207-2013, grant 258236) and FP7 collaborative project SYSCOL and COST Action in the UK is also acknowledged (BM1206). The COIN and COIN-B trials were funded by Cancer Research UK and the Medical Research Council and were conducted with the support of the National Institute of Health Research Cancer Research Network. COIN and COIN-B translational studies were supported by the Bobby Moore Fund from Cancer Research UK, Tenovus, the Kidani Trust, Cancer Research Wales and the National Institute for Social Care and Health Research Cancer Genetics Biomedical Research Unit (2011–2014). N.A.A., B.F.M. and S.M.W. were funded and supported by KFSHRC. In Finland, this work was supported by grants from the Academy of Finland (Finnish Center of Excellence Program 2012–2017, 250345), the Jane and Aatos Erkko Foundation, the Finnish Cancer Society (K.P.), the European Research Council [ERC; 268648], the Sigrid Juselius Foundation, SYSCOL, the Nordic Information for Action eScience Center (NIASC), the Nordic Center of Excellence financed by NordForsk (project 62721, K.P.) and State Research Funding of Kuopio University Hospital (B1401). We acknowledge the computational resources provided by the ELIXIR node, hosted at the CSC–IT Center for Science, Finland, and funded by the Academy of Finland (grants 271642 and 263164), the Ministry of Education and Culture, Finland. V.S. was supported by the Finnish Academy (grant number 139635). Sample collection and genotyping in the Finnish Twin Cohort has been supported by the Wellcome Trust Sanger Institute, ENGAGE — European Network for Genetic and Genomic Epidemiology, FP7-HEALTH-F4-2007 (201413), the National Institute of Alcohol Abuse and Alcoholism [grants AA-12502 and AA-00145 to Richard J Rose and K02AA018755 to Danielle M Dick] and the Academy of Finland (100499, 205585, 265240 and 263278 to J.K.). The work of the Colon Cancer Family Registry (CCFR) was supported by from the National Cancer Institute (UMI CA167551), National Institutes of Health and through cooperative agreements with the following CCFR centers: Australasian Colorectal Cancer Family Registry (U01 CA074778, U01/U24 CA097735), USC Consortium Colorectal Cancer Family Registry (U01/U24 CA074799), Mayo Clinic Cooperative Familial Registry for Colon Cancer Studies (U01/U24 CA074800), Ontario Familial Colorectal Cancer Registry (U01/U24 CA074783), Seattle Colorectal Cancer Family Registry (U01/U24 CA074794), and University of Hawaii Colorectal Cancer Family Registry (U01/U24 CA074806). The CCFR Illumina GWAS was supported by funding from the National Cancer Institute, National Institutes of Health (U01 CA122839 and R01 CA143237 to G.C.). Additional support was provided from the Surveillance, Epidemiology and End Results (SEER) Program of the National Cancer Institute to Fred Hutchinson Cancer Research Center (Control Nos. N01-CN-67009, N01-PC-35142 and Contract No. HHSN261201300012I), the Hawai’i Department of Health (Control Nos. N01-PC-67001, N01-PC-35137, and Contract No. HHSN26120100037C), and the California Department of Public Health (contract HHSN26120100035C) awarded to the University of Southern California. The content of this manuscript does not necessarily reflect the views or policies of the National Cancer Institute, any SEER program or any of the collaborating centres in the CCFR, nor does mention of trade names, commercial products, or organisations imply endorsement by the US Government, SEER or the CCFR.

We are grateful to all individuals who participated in the various studies. This study made use of genotyping data from the 1958 Birth Cohort, kindly made available by the Wellcome Trust Case Control Consortium 2. A full list of the investigators who contributed to the
generation of the data is available at http://www.wtccc.org.uk/.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.ejca.2017.07.034.

References


