Title: A combined proteomics and Mendelian randomization approach to investigate the effects of aspirin-targeted proteins on colorectal cancer

Authors:

Aayah Nounu[1,2], Alexander Greenhough[2,3], Kate J Heesom[4], Rebecca C Richmond[1], Jie Zheng[1], Stephanie J Weinstein[5], Demetrius Albanes[5], John A Baron[6], John L Hopper[7,8], Jane C Figueiredo[9,10], Polly A Newcomb[11,12], Noralane M Lindor[13], Graham Casey[14], Elizabeth A Platz[15], Loïc Le Marchand[16], Cornelia M Ulrich[17], Christopher I Li[18], Fränzel JB van Duijnhoven[19], Andrea Gsur[20], Peter T Campbell[21], Víctor Moreno[22,23,24,25], Pavel Vodicka[26,27,28], Ludmila Vodickova[26,27,28], Hermann Brenner[29,30,31], Jenny Chang-Claude[32,33], Michael Hoffmeister[29], Lori C Sakoda[18,34], Martha L Slattery[35], Robert E Schoen[36], Marc J Gunter[37], Sergi Castellví-Bel[38], Hyeong Rok Kim[39], Sun-Seog Kweon[40,41], Andrew T Chan[42,43,44,45,46,47], Li Li[48], Wei Zheng[49], D Timothy Bishop[50], Daniel D Buchanan[51,52,53], Graham G Giles[54,55], Stephen B Gruber[56], Gad Rennert[57,58,59], Zsofia K Stadler[60], Tabitha A Harrison[18], Yi Lin[18], Temitope O Keku[61], Michael O Woods[62], Clemens Schafmayer[63], Bethany Van Guelpen[64,65], Steven J Gallinger[66], Heather Hampel[67], Sonja I Berndt[5], Paul D P Pharoah[68], Annika Lindblom[69,70], Alicja Wolk[71,72], Anna H Wu[73], Emily White[18,74], Ulrike Peters[18,74], David A Drew[75], Dominique Scherer[76], Justo Lorenzo Bermejo[76], Ann C Williams[2], Caroline L Relton[1]

[1] Medical Research Council (MRC) Integrative Epidemiology Unit, Bristol Medical School, University of Bristol, Bristol, BS8 2BN, UK

[2]School of Cellular and Molecular Medicine, University of Bristol, Bristol, BS8 1TD, UK

[3]Centre for Research in Biosciences, The Faculty of Health and Applied Sciences, The University of the West of England, Bristol, BS16 1QY, UK

[4]Proteomics Facility, Faculty of Life Sciences, University of Bristol, Bristol, UK

[5]Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Bethesda, Maryland, USA.

[6]Department of Medicine, University of North Carolina School of Medicine, Chapel Hill, North Carolina, USA.

[7] Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global Health, The University of Melbourne, Melbourne, Victoria, Australia.

[8] Department of Epidemiology, School of Public Health and Institute of Health and Environment, Seoul National University, Seoul, South Korea.

[9] Department of Medicine, Samuel Oschin Comprehensive Cancer Institute, Cedars-Sinai Medical Center, Los Angeles, CA, USA.

[10] Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, California, USA.

[11] Public Health Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, Washington, USA

[12]School of Public Health, University of Washington, Seattle, Washington, USA

[13] Department of Health Science Research, Mayo Clinic, Scottsdale, Arizona, USA

[14] Center for Public Health Genomics, University of Virginia, Charlottesville, Virginia, USA.

[15] Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, USA

[16] University of Hawaii Cancer Center, Honolulu, Hawaii, USA

[17] Huntsman Cancer Institute and Department of Population Health Sciences, University of Utah, Salt Lake City, Utah, USA.

[18] Public Health Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, Washington, USA.

[19] Division of Human Nutrition and Health, Wageningen University & Research, Wageningen, The Netherlands.

[20] Institute of Cancer Research, Department of Medicine I, Medical University Vienna, Vienna, Austria.

[21] Behavioral and Epidemiology Research Group, American Cancer Society, Atlanta, Georgia, USA.

[22] Oncology Data Analytics Program, Catalan Institute of Oncology-IDIBELL, L'Hospitalet de Llobregat, Barcelona, Spain.

[23] CIBER Epidemiología y Salud Pública (CIBERESP), Madrid, Spain.

[24] Department of Clinical Sciences, Faculty of Medicine, University of Barcelona, Barcelona, Spain.

[25] ONCOBEL Program, Bellvitge Biomedical Research Institute (IDIBELL), L'Hospitalet de Llobregat, Barcelona, Spain.

[26] Department of Molecular Biology of Cancer, Institute of Experimental Medicine of the Czech Academy of Sciences, Prague, Czech Republic.

[27] Institute of Biology and Medical Genetics, First Faculty of Medicine, Charles University, Prague, Czech Republic.

[28] Faculty of Medicine and Biomedical Center in Pilsen, Charles University, Pilsen, Czech Republic.

[29] Division of Clinical Epidemiology and Aging Research, German Cancer Research Center (DKFZ), Heidelberg, Germany.

[30] Division of Preventive Oncology, German Cancer Research Center (DKFZ) and National Center for Tumor Diseases (NCT), Heidelberg, Germany.

[31] German Cancer Consortium (DKTK), German Cancer Research Center (DKFZ), Heidelberg, Germany.

[32] Division of Cancer Epidemiology, German Cancer Research Center (DKFZ), Heidelberg, Germany.

[33] University Medical Centre Hamburg-Eppendorf, University Cancer Centre Hamburg (UCCH), Hamburg, Germany.

[34] Division of Research, Kaiser Permanente Northern California, Oakland, California, USA.

[35] Department of Internal Medicine, University of Utah, Salt Lake City, Utah, USA.

[36] Department of Medicine and Epidemiology, University of Pittsburgh Medical Center, Pittsburgh, Pennsylvania, USA.

[37] Nutrition and Metabolism Section, International Agency for Research on Cancer, World Health Organization, Lyon, France.

[38] Gastroenterology Department, Hospital Clínic, Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Centro de Investigación Biomédica en Red de Enfermedades Hepáticas y Digestivas (CIBEREHD), University of Barcelona, Barcelona, Spain.

[39] Department of Surgery, Chonnam National University Hwasun Hospital and Medical School, Hwasun, Korea.

[40] Department of Preventive Medicine, Chonnam National University Medical School, Gwangju, Korea.

[41] Jeonnam Regional Cancer Center, Chonnam National University Hwasun Hospital, Hwasun, Korea.

[42] Division of Gastroenterology, Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts, USA.

[43] Channing Division of Network Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts, USA.

[44] Clinical and Translational Epidemiology Unit, Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts, USA.

[45] Broad Institute of Harvard and MIT, Cambridge, Massachusetts, USA.

[46] Department of Epidemiology, Harvard T.H. Chan School of Public Health, Harvard University, Boston, Massachusetts, USA.

[47] Department of Immunology and Infectious Diseases, Harvard T.H. Chan School of Public Health, Harvard University, Boston, Massachusetts, USA.

[48] Department of Family Medicine, University of Virginia, Charlottesville, Virginia, USA.

[49] Division of Epidemiology, Department of Medicine, Vanderbilt-Ingram Cancer Center, Vanderbilt Epidemiology Center, Vanderbilt University School of Medicine, Nashville, Tennessee, USA.

[50] Leeds Institute of Cancer and Pathology, University of Leeds, Leeds, UK.

[51] Colorectal Oncogenomics Group, Department of Clinical Pathology, The University of Melbourne, Parkville, Victoria 3010 Australia

[52] University of Melbourne Centre for Cancer Research, Victorian Comprehensive Cancer Centre, Parkville, Victoria 3010 Australia

[53] Genetic Medicine and Family Cancer Clinic, The Royal Melbourne Hospital, Parkville, Victoria, Australia.

[54] Cancer Epidemiology Division, Cancer Council Victoria, Melbourne, Victoria, Australia.

[55] Precision Medicine, School of Clinical Sciences at Monash Health, Monash University, Clayton, Victoria, Australia.

[56] Department of Preventive Medicine & USC Norris Comprehensive Cancer Center, Keck School of Medicine, University of Southern California, Los Angeles, California, USA.

[57] Department of Community Medicine and Epidemiology, Lady Davis Carmel Medical Center, Haifa, Israel.

[58] Ruth and Bruce Rappaport Faculty of Medicine, Technion-Israel Institute of Technology, Haifa, Israel.

[59] Clalit National Cancer Control Center, Haifa, Israel.

[60] Department of Medicine, Memorial Sloan Kettering Cancer Center, New York, New York, USA.

[61] Center for Gastrointestinal Biology and Disease, University of North Carolina, Chapel Hill, North Carolina, USA.

[62] Memorial University of Newfoundland, Discipline of Genetics, St. John's, Canada.

[63] Department of General Surgery, University Hospital Rostock, Rostock, Germany.

[64] Department of Radiation Sciences, Oncology Unit, Umeå University, Umeå, Sweden.

[65] Wallenberg Centre for Molecular Medicine, Umeå University, Umeå, Sweden.

[66] Lunenfeld Tanenbaum Research Institute, Mount Sinai Hospital, University of Toronto, Toronto, Ontario, Canada.

[67] Division of Human Genetics, Department of Internal Medicine, The Ohio State University Comprehensive Cancer Center, Columbus, Ohio, USA.

[68] Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK.

[69] Department of Clinical Genetics, Karolinska University Hospital, Stockholm, Sweden.

[70] Department of Molecular Medicine and Surgery, Karolinska Institutet, Stockholm, Sweden.

[71] Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden.

[72] Department of Surgical Sciences, Uppsala University, Uppsala, Sweden.

[73] University of Southern California, Preventative Medicine, Los Angeles, California, USA.

[74] Department of Epidemiology, University of Washington School of Public Health, Seattle, Washington, USA.

[75] Massachusetts General Hospital and Harvard Medical School, Clinical and Translational Epidemiology Unit, Boston, Massachusetts 02114, USA.

[76] Institute of Medical Biometry and Informatics, University of Heidelberg, Im Neuenheimer Feld 130.3, Heidelberg, Germany.

Running title: Aspirin-targeted proteins and risk of colorectal cancer

Keywords: aspirin, proteome, Mendelian randomization, colorectal

Author funding

This work was supported funded by a PhD studentship from the Medical Research Council (AN), a Cancer Research UK Programme Grant (C19/A11975, TJC, ACW), an MRC Research Grant (MR/R017247/1, TJC, ACW) and by the John James Bristol Foundation. Further funding was provided by The UK Medical Research Council Integrative Epidemiology Unit (MC_UU_12013_2, CLR) and Cancer Research UK (C18281/A19169, CLR), the Wellcome Trust (217487/Z/19/Z, AG) and Bowel Cancer UK (19PT0039 AG). RCR is a de Pass Vice Chancellor Research Fellow at the University of Bristol.

Consortia funding

Genetics and Epidemiology of Colorectal Cancer Consortium (GECCO): National Cancer Institute, National Institutes of Health, U.S. Department of Health and Human Services (U01 CA164930, U01 CA137088, R01 CA059045, R21 CA191312, R01201407). Genotyping/Sequencing services were provided by the Center for Inherited Disease Research (CIDR) (X01-HG008596 and X-01-HG007585). CIDR is fully funded through a federal contract from the National Institutes of Health to The Johns Hopkins University, contract number HHSN2682012000081. This research was funded in part through the NIH/NCI Cancer Center Support Grant P30 CA015704.

ASTERISK: a Hospital Clinical Research Program (PHRC-BRD09/C) from the University Hospital Center of Nantes (CHU de Nantes) and supported by the Regional Council of Pays de la Loire, the Groupement des Entreprises Françaises dans la Lutte contre le Cancer (GEFLUC), the Association Anne de Bretagne Génétique and the Ligue Régionale Contre le Cancer (LRCC).

The ATBC Study is supported by the Intramural Research Program of the U.S. National Cancer Institute, National Institutes of Health.

CLUE funding was from the National Cancer Institute (U01 CA86308, Early Detection Research Network; P30 CA006973), National Institute on Aging (U01 AG18033), and the American Institute for Cancer Research. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial

products, or organizations imply endorsement by the US government.COLO2&3: National Institutes of Health (R01 CA60987).

ColoCare: This work was supported by the National Institutes of Health (grant numbers R01 CA189184 (Li/Ulrich), U01 CA206110 (Ulrich/Li/Siegel/Figueireido/Colditz, 2P30CA015704- 40 (Gilliland), R01 CA207371 (Ulrich/Li)), the Matthias Lackas-Foundation, the German Consortium for Translational Cancer Research, and the EU TRANSCAN initiative.

The Colon Cancer Family Registry (CCFR, www.coloncfr.org) is supported in part by funding from the National Cancer Institute (NCI), National Institutes of Health (NIH) (award U01 CA167551). The CCFR Set-1 (Illumina 1M/1M-Duo) and Set-2 (Illumina Omni1-Quad) scans were supported by NIH awards U01 CA122839 and R01 CA143247 (to GC). The CCFR Set-3 (Affymetrix Axiom CORECT Set array) was supported by NIH award U19 CA148107 and R01 CA81488 (to SBG). The CCFR Set-4 (Illumina OncoArray 600K SNP array) was supported by NIH award U19 CA148107 (to SBG) and by the Center for Inherited Disease Research (CIDR), which is funded by the NIH to the Johns Hopkins University, contract number HHSN268201200008I. The SCCFR Illumina HumanCytoSNP array was supported through NCI award R01 CA076366 (to PAN). Additional funding for the OFCCR/ARCTIC was through award GL201-043 from the Ontario Research Fund (to BWZ), award 112746 from the Canadian Institutes of Health Research (to TJH), through a Cancer Risk Evaluation (CaRE) Program grant from the Canadian Cancer Society (to SG), and through generous support from the Ontario Ministry of Research and Innovation. The content of this manuscript does not necessarily reflect the views or policies of the NCI, NIH or any of the collaborating centers in the Colon Cancer Family Registry (CCFR), nor does mention of trade names, commercial products, or organizations imply endorsement by the US Government, any cancer registry, or the CCFR.

COLON: The COLON study is sponsored by Wereld Kanker Onderzoek Fonds, including funds from grant 2014/1179 as part of the World Cancer Research Fund International Regular Grant Programme, by Alpe d'Huzes and the Dutch Cancer Society (UM 2012–5653, UW 2013-5927, UW2015-7946), and by TRANSCAN (JTC2012-MetaboCCC, JTC2013-FOCUS). The Nqplus study is

197 sponsored by a ZonMW investment grant (98-10030); by PREVIEW, the project PREVention of
198 diabetes through lifestyle intervention and population studies in Europe and around the World
199 (PREVIEW) project which received funding from the European Union Seventh Framework
200 Programme (FP7/2007–2013) under grant no. 312057; by funds from TI Food and Nutrition
201 (cardiovascular health theme), a public–private partnership on precompetitive research in food and
202 nutrition; and by FOODBALL, the Food Biomarker Alliance, a project from JPI Healthy Diet for a
203 Healthy Life.

204 Colorectal Cancer Transdisciplinary (CORECT) Study: The CORECT Study was supported by the
205 National Cancer Institute, National Institutes of Health (NCI/NIH), U.S. Department of Health and
206 Human Services (grant numbers U19 CA148107, R01 CA81488, P30 CA014089, R01 CA197350,; P01
207 CA196569; R01 CA201407) and National Institutes of Environmental Health Sciences, National
208 Institutes of Health (grant number T32 ES013678).

209 CORSA: “Österreichische Nationalbank Jubiläumsfondsprojekt” (12511) and Austrian Research
210 Funding Agency (FFG) grant 829675.

211 CPS-II: The American Cancer Society funds the creation, maintenance, and updating of the Cancer
212 Prevention Study-II (CPS-II) cohort. This study was conducted with Institutional Review Board
213 approval.

214 CRCGEN: Colorectal Cancer Genetics & Genomics, Spanish study was supported by Instituto de Salud
215 Carlos III, co-funded by FEDER funds –a way to build Europe– (grants PI14-613 and PI09-1286),
216 Agency for Management of University and Research Grants (AGAUR) of the Catalan Government
217 (grant 2017SGR723), and Junta de Castilla y León (grant LE22A10-2). Sample collection of this work
218 was supported by the Xarxa de Bancs de Tumors de Catalunya sponsored by Pla Director d’Oncología
219 de Catalunya (XBTC), Plataforma Biobancos PT13/0010/0013 and ICOBIOBANC, sponsored by the
220 Catalan Institute of Oncology.

Czech Republic CCS: This work was supported by the Grant Agency of the Czech Republic (grants CZ GA CR: GAP304/10/1286 and 1585) and by the Grant Agency of the Ministry of Health of the Czech Republic (grants AZV 15-27580A and AZV 17-30920A).

DACHS: This work was supported by the German Research Council (BR 1704/6-1, BR 1704/6-3, BR 1704/6-4, CH 117/1-1, HO 5117/2-1, HE 5998/2-1, KL 2354/3-1, RO 2270/8-1 and BR 1704/17-1), the Interdisciplinary Research Program of the National Center for Tumor Diseases (NCT), Germany, and the German Federal Ministry of Education and Research (01KH0404, 01ER0814, 01ER0815, 01ER1505A and 01ER1505B).

DALS: National Institutes of Health (R01 CA48998 to M. L. Slattery).

EDRN: This work is funded and supported by the NCI, EDRN Grant (U01 CA 84968-06).

EPIC: The coordination of EPIC is financially supported by the European Commission (DGSANCO) and the International Agency for Research on Cancer. The national cohorts are supported by Danish Cancer Society (Denmark); Ligue Contre le Cancer, Institut Gustave Roussy, Mutuelle Générale de l'Education Nationale, Institut National de la Santé et de la Recherche Médicale (INSERM) (France); German Cancer Aid, German Cancer Research Center (DKFZ), Federal Ministry of Education and Research (BMBF), Deutsche Krebshilfe, Deutsches Krebsforschungszentrum and Federal Ministry of Education and Research (Germany); the Hellenic Health Foundation (Greece); Associazione Italiana per la Ricerca sul Cancro-AIRCItaly and National Research Council (Italy); Dutch Ministry of Public Health, Welfare and Sports (VWS), Netherlands Cancer Registry (NKR), LK Research Funds, Dutch Prevention Funds, Dutch ZON (Zorg Onderzoek Nederland), World Cancer Research Fund (WCRF), Statistics Netherlands (The Netherlands); ERC-2009-AdG 232997 and Nordforsk, Nordic Centre of Excellence programme on Food, Nutrition and Health (Norway); Health Research Fund (FIS), PI13/00061 to Granada, PI13/01162 to EPIC-Murcia, Regional Governments of Andalucía, Asturias, Basque Country, Murcia and Navarra, ISCIII RETIC (RD06/0020) (Spain); Swedish Cancer Society, Swedish Research Council and County Councils of Skåne and Västerbotten (Sweden); Cancer

246 Research UK (14136 to EPIC-Norfolk; C570/A16491 and C8221/A19170 to EPIC-Oxford), Medical
247 Research Council (1000143 to EPIC-Norfolk, MR/M012190/1 to EPICOxford) (United Kingdom).

248 EPICOLON: This work was supported by grants from Fondo de Investigación Sanitaria/FEDER
249 (PI08/0024, PI08/1276, PS09/02368, PI11/00219, PI11/00681, PI14/00173, PI14/00230, PI17/00509,
250 17/00878, Acción Transversal de Cáncer), Xunta de Galicia (PGIDIT07PXIB9101209PR), Ministerio de
251 Economia y Competitividad (SAF07-64873, SAF 2010-19273, SAF2014-54453R), Fundación Científica
252 de la Asociación Española contra el Cáncer (GCB13131592CAST), Beca Grupo de Trabajo "Oncología"
253 AEG (Asociación Española de Gastroenterología), Fundación Privada Olga Torres, FP7 CHIBCHA
254 Consortium, Agència de Gestió d'Ajuts Universitaris i de Recerca (AGAUR, Generalitat de Catalunya,
255 2014SGR135, 2014SGR255, 2017SGR21, 2017SGR653), Catalan Tumour Bank Network (Pla Director
256 d'Oncologia, Generalitat de Catalunya), PERIS (SLT002/16/00398, Generalitat de Catalunya), CERCA
257 Programme (Generalitat de Catalunya) and COST Action BM1206 and CA17118. CIBERehd is funded
258 by the Instituto de Salud Carlos III.

259 ESTHER/VERDI. This work was supported by grants from the Baden-Württemberg Ministry of
260 Science, Research and Arts and the German Cancer Aid.

261 Harvard cohorts (HPFS, NHS, PHS): HPFS is supported by the National Institutes of Health (P01
262 CA055075, UM1 CA167552, U01 CA167552, R01 CA137178, R01 CA151993, R35 CA197735, K07
263 CA190673, and P50 CA127003), NHS by the National Institutes of Health (R01 CA137178, P01
264 CA087969, UM1 CA186107, R01 CA151993, R35 CA197735, K07CA190673, and P50 CA127003) and
265 PHS by the National Institutes of Health (R01 CA042182).

266 Hawaii Adenoma Study: NCI grants R01 CA72520.

267 HCES-CRC: the Hwasun Cancer Epidemiology Study–Colon and Rectum Cancer (HCES-CRC; grants
268 from Chonnam National University Hwasun Hospital, HCRI15011-1).

269 Kentucky: This work was supported by the following grant support: Clinical Investigator Award from
270 Damon Runyon Cancer Research Foundation (CI-8); NCI R01CA136726.

271 LCCS: The Leeds Colorectal Cancer Study was funded by the Food Standards Agency and Cancer
272 Research UK Programme Award (C588/A19167).

273 MCCS cohort recruitment was funded by VicHealth and Cancer Council Victoria. The MCCS was
274 further supported by Australian NHMRC grants 509348, 209057, 251553 and 504711 and by
275 infrastructure provided by Cancer Council Victoria. Cases and their vital status were ascertained
276 through the Victorian Cancer Registry (VCR) and the Australian Institute of Health and Welfare
277 (AIHW), including the National Death Index and the Australian Cancer Database.

278 Multiethnic Cohort (MEC) Study: National Institutes of Health (R37 CA54281, P01 CA033619, R01
279 CA063464 and U01 CA164973).

280 MECC: This work was supported by the National Institutes of Health, U.S. Department of Health and
281 Human Services (R01 CA81488 to SBG and GR).

282 MSKCC: The work at Sloan Kettering in New York was supported by the Robert and Kate Niehaus
283 Center for Inherited Cancer Genomics and the Romeo Milio Foundation. Moffitt: This work was
284 supported by funding from the National Institutes of Health (grant numbers R01 CA189184, P30
285 CA076292), Florida Department of Health Bankhead-Coley Grant 09BN-13, and the University of
286 South Florida Oehler Foundation. Moffitt contributions were supported in part by the Total Cancer
287 Care Initiative, Collaborative Data Services Core, and Tissue Core at the H. Lee Moffitt Cancer Center
288 & Research Institute, a National Cancer Institute-designated Comprehensive Cancer Center (grant
289 number P30 CA076292).

290 NCCCS I & II: We acknowledge funding support for this project from the National Institutes of Health,
291 R01 CA66635 and P30 DK034987.

292 NFCCR: This work was supported by an Interdisciplinary Health Research Team award from the
293 Canadian Institutes of Health Research (CRT 43821); the National Institutes of Health, U.S.
294 Department of Health and Human Serivces (U01 CA74783); and National Cancer Institute of Canada
295 grants (18223 and 18226). The authors wish to acknowledge the contribution of Alexandre Belisle

296 and the genotyping team of the McGill University and Génome Québec Innovation Centre, Montréal,
297 Canada, for genotyping the Sequenom panel in the NFCCR samples. Funding was provided to
298 Michael O. Woods by the Canadian Cancer Society Research Institute.

299 NSHDS: Swedish Research Council; Swedish Cancer Society; Cutting-Edge Research Grant and other
300 grants from Region Västerbotten; Knut and Alice Wallenberg Foundation; Lion's Cancer Research
301 Foundation at Umeå University; the Cancer Research Foundation in Northern Sweden; and the
302 Faculty of Medicine, Umeå University, Umeå, Sweden.

303 OFCCR: The Ontario Familial Colorectal Cancer Registry was supported in part by the National Cancer
304 Institute (NCI) of the National Institutes of Health (NIH) under award U01 CA167551 and award
305 U01/U24 CA074783 (to SG). Additional funding for the OFCCR and ARCTIC testing and genetic
306 analysis was through and a Canadian Cancer Society CaRE (Cancer Risk Evaluation) program grant
307 and Ontario Research Fund award GL201-043 (to BWZ), through the Canadian Institutes of Health
308 Research award 112746 (to TJH), and through generous support from the Ontario Ministry of
309 Research and Innovation.OSUMC: OCCPI funding was provided by Pelotonia and HNPCC funding was
310 provided by the NCI (CA16058 and CA67941).

311 PLCO: Intramural Research Program of the Division of Cancer Epidemiology and Genetics and
312 supported by contracts from the Division of Cancer Prevention, National Cancer Institute, NIH,
313 DHHS. Funding was provided by National Institutes of Health (NIH), Genes, Environment and Health
314 Initiative (GEI) Z01 CP 010200, NIH U01 HG004446, and NIH GEI U01 HG 004438.

315 SCCFR: The Seattle Colon Cancer Family Registry was supported in part by the National Cancer
316 Institute (NCI) of the National Institutes of Health (NIH) under awards U01 CA167551, U01 CA074794
317 (to JDP), and awards U24 CA074794 and R01 CA076366 (to PAN).

318 SEARCH: The University of Cambridge has received salary support in respect of PDPP from the NHS in
319 the East of England through the Clinical Academic Reserve. Cancer Research UK (C490/A16561); the
320 UK National Institute for Health Research Biomedical Research Centres at the University of
321 Cambridge.

SELECT: Research reported in this publication was supported in part by the National Cancer Institute of the National Institutes of Health under Award Numbers U10 CA37429 (CD Blanke), and UM1 CA182883 (CM Tangen/IM Thompson). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

SMS: This work was supported by the National Cancer Institute (grant P01 CA074184 to J.D.P. and P.A.N., grants R01 CA097325, R03 CA153323, and K05 CA152715 to P.A.N., and the National Center for Advancing Translational Sciences at the National Institutes of Health (grant KL2 TR000421 to A.N.B.-H.)

The Swedish Low-risk Colorectal Cancer Study: The study was supported by grants from the Swedish research council; K2015-55X-22674-01-4, K2008-55X-20157-03-3, K2006-72X-20157-01-2 and the Stockholm County Council (ALF project).

Swedish Mammography Cohort and Cohort of Swedish Men: This work is supported by the Swedish Research Council /Infrastructure grant, the Swedish Cancer Foundation, and the Karolinska Institute´s Distinguished Professor Award to Alicja Wolk.

UK Biobank: This research has been conducted using the UK Biobank Resource under Application Number 8614

VITAL: National Institutes of Health (K05 CA154337).

WHI: The WHI program is funded by the National Heart, Lung, and Blood Institute, National Institutes of Health, U.S. Department of Health and Human Services through contracts HHSN268201100046C, HHSN268201100001C, HHSN268201100002C, HHSN268201100003C, HHSN268201100004C, and HHSN271201100004C.

Corresponding author:

Name: Aayah Nounu

E-mail address an0435@bristol.ac.uk

346 Address: Oakfield House, Oakfield Grove, Clifton, Bristol, UK, BS8 2BN

347

348 Conflict of interest

349 Elizabeth A. Platz is the Editor-in-Chief of Caner Epidemiology, Biomarkers and Prevention.

350

351 Manuscript word count: 3,918

352 Number of figures: 2

353 Number of Tables: 1

354

355 Abstract

356 Background: Evidence for aspirin's chemopreventative properties on colorectal cancer (CRC) is
357 substantial, but its mechanism of action is not well-understood. We combined a proteomic approach
358 with Mendelian randomization (MR) to identify possible new aspirin targets that decrease CRC risk.

359 Methods: Human colorectal adenoma cells (RG/C2) were treated with aspirin (24 hours) and a stable
360 isotope labelling with amino acids in cell culture (SILAC) based proteomics approach identified
361 altered protein expression. Protein quantitative trait loci (pQTLs) from INTERVAL (N=3,301) and
362 expression QTLs (eQTLs) from the eQTLGen Consortium (N=31,684) were used as genetic proxies for
363 protein and mRNA expression levels. Two-sample MR of mRNA/protein expression on CRC risk was
364 performed using eQTL/pQTL data combined with CRC genetic summary data from the Colon Cancer
365 Family Registry (CCFR), Colorectal Transdisciplinary (CORECT), Genetics and Epidemiology of
366 Colorectal Cancer (GECCO) consortia and UK Biobank (55,168 cases and 65,160 controls).

367 Results: Altered expression was detected for 125/5886 proteins. Of these, aspirin decreased MCM6,
368 RRM2 and ARFIP2 expression and MR analysis showed that a standard deviation increase in
369 mRNA/protein expression was associated with increased CRC risk (OR:1.08, 95% CI:1.03-1.13,
370 OR:3.33, 95% CI:2.46-4.50 and OR:1.15, 95% CI:1.02-1.29, respectively).

371 Conclusion: MCM6 and RRM2 are involved in DNA repair whereby reduced expression may lead to
372 increased DNA aberrations and ultimately cancer cell death, whereas ARFIP2 is involved in actin
373 cytoskeletal regulation indicating a possible role in aspirin's reduction of metastasis.

374 Impact: Our approach has shown how laboratory experiments and population-based approaches can
375 combine to identify aspirin-targeted proteins possibly affecting CRC risk.

376

Introduction

Colorectal cancer (CRC) is the fourth most common cancer worldwide (1). Observational studies as well as randomized controlled trials (RCTs) using aspirin for the prevention of vascular events have shown that aspirin use is associated with a decrease in CRC incidence and mortality (2–5). This was primarily thought to be through the acetylation of the cyclooxygenase (COX) enzymes thereby inhibiting their action (6). These enzymes are involved in the COX/prostaglandin E2(PGE_2) signalling pathway which is frequently upregulated in CRC, driving many of the hallmarks of cancer (7,8).

Evidence for COX-independent mechanisms have also emerged, such as the prevention of NFκB activation, inhibition of the extracellular-signal-regulated kinase (ERK) signalling pathway, cell cycle progression inhibition and possible induction of autophagy (7,9). An aspirin derivative that does not inhibit COX reduced the mean number of aberrant crypt foci (an early lesion in colorectal carcinogenesis) in a mouse model of CRC more than aspirin itself (10). Furthermore, aspirin was able to inhibit proliferation and induce apoptosis in COX-2 negative colon cancer cell lines as well as reducing angiogenesis in 3D assays where COX-inhibitors showed no effect (11–13). Clinically, aspirin has been shown to reduce tumour recurrence in phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit alpha (PIK3CA) mutant cancer whereas rofecoxib (a COX-2 selective inhibitor) showed no effect (14) and has also been shown to improve survival in patients with human leukocyte antigen (HLA) class I antigen expression, regardless of COX-2 expression (15). There is now a significant number of studies that indicate the mechanism behind the action of aspirin on CRC risk is still not fully understood and that multiple mechanisms are involved (16).

In conventional epidemiological studies it is often difficult to determine causality due to limitations of confounding and reverse causation. While RCTs can overcome these limitations, they are generally limited to assessing the causal role of health interventions or pharmaceutical agents on disease outcomes, rather than understanding biological mechanisms. Furthermore, in the context of cancer, RCTs for cancer primary prevention are not always feasible, as they require long-term follow-up for the cancer to develop. Mendelian randomization (MR) is an epidemiological method which

403 applies a similar notion of randomization as in the RCT to evaluate causality. In MR, genetic variants
404 (most commonly single nucleotide polymorphisms (SNPs)) are used to proxy an exposure of interest
405 (17). As genetic variants are randomly assorted at conception, an individual's genetic makeup is
406 unlikely to be influenced by exposures later on in life, thus reducing the possibility of confounding
407 and reverse causation (18).

408 More recently, the increase in genome-wide association studies for molecular traits has identified
409 SNPs that are associated with protein and mRNA expression levels, thereby providing protein
410 quantitative trait loci (pQTLs) and expression quantitative trait loci (eQTLs) (19,20), which may be
411 used to investigate the causal mechanism of drug targets on disease risk (21).Such methods can
412 complement laboratory experiments to better understand the mechanism of action of drugs on
413 cancer growth and progression.

414 Due to evidence showing that aspirin may prevent adenoma formation (22) and adenomas being the
415 precursors of most colorectal cancers (23), we focused on a colorectal adenoma cell line (RG/C2) in
416 this study and identified altered protein expression in relation to aspirin treatment. Findings were
417 then taken forward into an MR analysis to investigate which proteins targeted by aspirin may be
418 causally implicated in reducing risk of CRC incidence, thereby providing insight into alternative
419 mechanisms/pathways for the action of aspirin.

420

Methods

Cell culture experiments

The S/RG/C2 (referred to as RG/C2 henceforth whereby the prefix "S" denotes that they are from a sporadic tumour) (RRID:CVCL_IQ11) colorectal adenoma cell line was derived in the Colorectal Tumour Biology group and is described in detail elsewhere (24). These cells express RG/C2 cells express WT full length *APC* (251) as well as wild type *KRAS* and *PIK3CA* (252) but express mutant *TP53* (25–27) .RG/C2s were cultured in Dulbecco's Modified Eagles Medium (DMEM) (Life Technologies, Paisley, UK) and supplemented with 20% foetal bovine serum (FBS)(Life Technologies, Paisley, UK), L-glutamine (2mM)(Life Technologies, Paisley, UK), penicillin (100 units/ml) (Life Technologies, Paisley, UK), streptomycin (100 ug/ml) (Life Technologies, Paisley, UK) and insulin (0.2 units/ml) (Sigma-Aldrich, Poole, UK). Cells were mycoplasma tested (Mycoalert Plus mycoplasma detection kit; Lonza Group, Basal, Switzerland) and experiments performed within 10 passages. Aspirin (Sigma-Aldrich) was dissolved in fresh growth medium and diluted to form concentrations of 2mM and 4mM.

Generation of proteomic data - SILAC approach

A stable isotope labelling with amino acids in cell culture (SILAC) approach was carried out on RG/C2 cells treated with 0mM, 2mM and 4mM aspirin for 24 hours. Control cells (0mM aspirin) were cultured with an L-arginine and L-lysine (light labelling), 2mM were cultured with $^{2}H_{4}$-lysine and $_{13}C_{6}$-arginine (medium labelling) and 4mM were cultured with $^{15}N_{2}{}^{13}C_{6}$-lysine and $^{15}N_{4}{}^{13}C_{6}$-arginine (heavy labelling) (Cambridge Isotope Laboratory, Massachusetts, United States). These methods were based on the SILAC-based mass spectrometry approach by Trinkle-Mulcahy et. al (2008) (28).

Cells were cultured with aspirin and the isotopes for 24 hours before extracting protein lysates. This experiment was carried out in duplicate. Lysates from the three conditions were pooled in a 1:1:1 ratio, separated by SDS-PAGE and then subjected to in-gel tryptic digestion. The resulting peptides were analysed by liquid chromatography mass spectrometry using an LTQ Orbitrap Velos mass spectrometer (Thermo Fisher Scientific, Waltham, Massachusetts, USA) and the mass spectral data

analysed using Proteome Discoverer software v1.4 (Thermo). Details of SILAC labelling and proteomics have been previously published (29). To determine proteins whose expression is altered due to aspirin treatment, we applied a threshold of a 1.4 fold change between 4mM/control and 2mM/control, as suggested previously (30). Results were also limited to a variability of <100% and a peptide count of at least 2.

Statistical analyses

Two-sample MR

To assess the effect of protein/mRNA expression of aspirin targets on risk of CRC, we used a two-sample MR approach. Firstly, SNPs were identified to proxy for protein/mRNA expression of the proteins shown to be altered in cell culture. Genetic association estimates with protein/mRNA expression levels (pQTLs/eQTLs) (sample 1) were integrated with genetic association estimates with CRC risk (sample 2).

Genetic predictors for protein and gene expression

Protein quantitative trait loci (pQTLs) were obtained from the INTERVAL study which comprises about 50,000 individuals within a randomised trial evaluating the effect of varying intervals between blood donations and how this affects outcomes such as quality of life (31). Relative protein measurements were taken using SOMAscan assays for 3,622 plasma proteins in a subset of 3,301 participants, randomly chosen. Genotyping and imputation (using a combined 1000 Genomes Phase 3-UK10K as the reference panel) of these individuals provided measures for 10,572,814 variants that passed quality control and were taken forward in a GWAS analysis to identify pQTLs for the measured proteins (details of quality control are mentioned elsewhere (19)). pQTLs identified represent a standard deviation (SD) change in protein expression (19).To adjust for multiple testing, a Bonferroni correction ($0.05/10{,}572{,}814=4.72\times10^{-9}$) was applied and pQTLs below this P-value threshold were used to proxy for protein expression in our analysis (32).

In the absence of a relevant pQTL for the protein of interest, an equivalent mRNA expression GWAS was used instead. Expression quantitative trait loci (eQLTs) were extracted from the eQTLGEN

consortium consisting of 31,684 individuals from 37 datasets, of which 26,886 samples were from blood and 4798 from peripheral blood mononuclear cells (PBMCs). Due to the differing methods for genotyping between the studies, variants for each transcript ranged between 2,337-31,684 variants (20). For this reason, a Bonferroni correction threshold was adjusted depending on the number of variants measured for each transcript (0.05/number of variants) (32). eQTLs were standardized and meta-analysed through a Z-transformation, therefore eQTL effect sizes are reported as standard deviation (SD) changes (20).

In this analysis, both cis (within 1 Mb of the gene transcription start sit) and trans QTLs were used to proxy for expression. Once suitable pQTLs/eQTLs were identified, linkage disequilibiurm (LD) clumping at an R^2 of 0.001 was carried out to remove SNPs that are inherited together and so that only the SNP most strongly associated with the mRNA/protein expression within a 10,000kb window was used.

Genetic association for colorectal cancer

Genetic association summary statistics for CRC, comprising 55,168 colorectal cancer cases and 65,160 controls, were obtained from the Colon Cancer Family Registry (CCFR), Colorectal Transdisciplinary (CORECT) and Genetics and Epidemiology of Colorectal Cancer (GECCO) consortia and UK Biobank (33–35). Quality control procedures have been described elsewhere (33). Ethics were approved by respective institutional review boards.

Evaluating the association of mRNA/protein expression on colorectal cancer

Analyses were carried out in R version 3.2.3 using the MR-Base TwoSampleMR R package (github.com/MRCIEU/TwoSampleMR) (36), which allows the formatting, harmonisation and analysis of summary statistics. The package reassigns alleles so that the effect allele has a positive association with the exposure and so represents an increase in protein/mRNA expression. In turn, allele harmonization ensures that the same allele (that predicts increased expression) is the effect allele in the outcome dataset as well. In the case of palindromic SNPs (represented by either A/T or G/C on both the forward and reverse alleles) these were also harmonized where possible based on allele

frequencies. If allele frequencies for the effect allele and the other allele were similar, thus making harmonization difficult, these SNPs were dropped from the analysis (36).

Separate MR analyses were carried for cis and trans pQTLs as well as cis and trans eQTLs. For proteins with just one pQTL or eQTL, Wald ratios (SNP-outcome estimate ÷ SNP-exposure estimate) were calculated to give a causal estimate for risk of CRC per SD increase in mRNA/protein expression. Where more than one QTL was available as a proxy for the exposure (mRNA/protein levels), a weighted mean of the ratio estimates weighted by the inverse variance of the ratio estimates (inverse-variance weighted (IVW) method) was used (37).

When one genetic variant used to proxy for an exposure is invalid e.g. due to horizontal pleiotropy (where a genetic variant affects the outcome through an alternative exposure/pathway of interest) (17), then the estimator from the IVW method becomes biased (38). As a sensitivity analysis, alternative MR methods were used when more than 2 SNPs were available as instruments for mRNA/protein expression (MR Egger, simple mode, weighted mode, and weighted median) (36,39,40). Unlike the IVW method, the MR Egger method is not constrained to pass through an effect size of 0, thereby allowing the assessment of horizontal pleiotropy through the y intercept. (38,41). The weighted median approach is useful as it allows a consistent estimate even if 50% of the SNPs proxying protein/mRNA expression are invalid instruments (40) and the mode estimate also provides a consistent causal effect estimate even if the majority of the instruments are invalid, as the estimate depends on the largest number of similar instruments (39).

Results

Mendelian randomization of gene/protein expression and risk of colorectal cancer identified in aspirin treated human adenoma cells

In order to investigate the early changes that could reduce cancer risk, we investigated the proteome of aspirin treated adenoma derived cells to identify new targets of aspirin that may alter the risk of CRC by combining these proteomic results with an MR analysis. After applying a filtering threshold based on fold change and variability in expression, we identified 125 proteins whose

525 expression appeared to be regulated by aspirin treatment (Figure 1) (S1 Table), although 5 were
526 uncharacterised from mass spectrometry and therefore excluded from the analysis.

527 Of the 120 proteins, expression of 28 proteins was measured in the INTERVAL study, of which 12
528 proteins had pQTLs that were below the Bonferroni significance threshold (0.05/10,572,814 = 4.73
529 $x10^{-9}$). From these 12 proteins, cis pQTLs were available for 3 proteins and trans pQTLs for 10
530 proteins (S2 Table). In the absence of available pQTLs, eQTLs for the transcripts of the identified
531 proteins were used instead. Of the 108 proteins with no pQTLs available, expression of 89 mRNAs
532 were measured in the eQTLGen consortium, of which 77 proteins had eQTLs that were below the
533 Bonferroni significance threshold. From these 77 proteins, cis eQTLs were available for 71 proteins
534 and trans eQTLs were available for 37 proteins (S3 Table). In total, there were 318 unique SNPs
535 proxying for protein and mRNA expression, of which outcome summary statistics were available for
536 305 SNPs to test for association between 99 mRNA/proteins against risk of CRC.

537 Two-sample MR analysis using the Wald ratio or IVW method was conducted to test the effect of
538 increased mRNA/protein expression on the risk of CRC incidence using cis and trans pQTLs (S4 Table)
539 as well as cis and trans eQTLs (S5 Table). In total, 99 proteins were tested for association with CRC
540 incidence. To correct for multiple testing, a Bonferroni adjusted threshold of significance was applied
541 ($0.05/99= 5.05x10^{-4}$) but we also considered associations of a nominal significance (P value<0.05).
542 Overall, 1 protein with cis eQTLs and 2 with trans eQTLs were associated with CRC incidence at P<
543 $5.05x10^{-4}$ and a further 3 proteins with cis eQTLs, 1 with a trans eQTL and 1 instrumented by a trans
544 pQTL were associated with CRC incidence at a P value < 0.05.

545 Increased mRNA expression of Human Leukocyte Antigen A (*HLA-A*) and mini chromosome
546 maintenance 6 (*MCM6*) instrumented by cis eQTLs was found to be associated with an increased risk
547 of CRC incidence (OR 1.28, 95% CI:1.04-1.58, P value: 0.02 and OR 1.08, 95% CI: 1.03-1.13, P value:
548 $9.23x10^{-4}$ per SD increase in mRNA expression, respectively). An SD increase in mRNA expression of
549 fatty acid desaturase 2 (*FADS2*) and DNA polymerase delta subunit 2 (*POLD2*) instrumented by cis
550 eQTLs was associated with a decrease in risk of CRC incidence (OR 0.94, 95% CI: 0.90-0.97, P value:

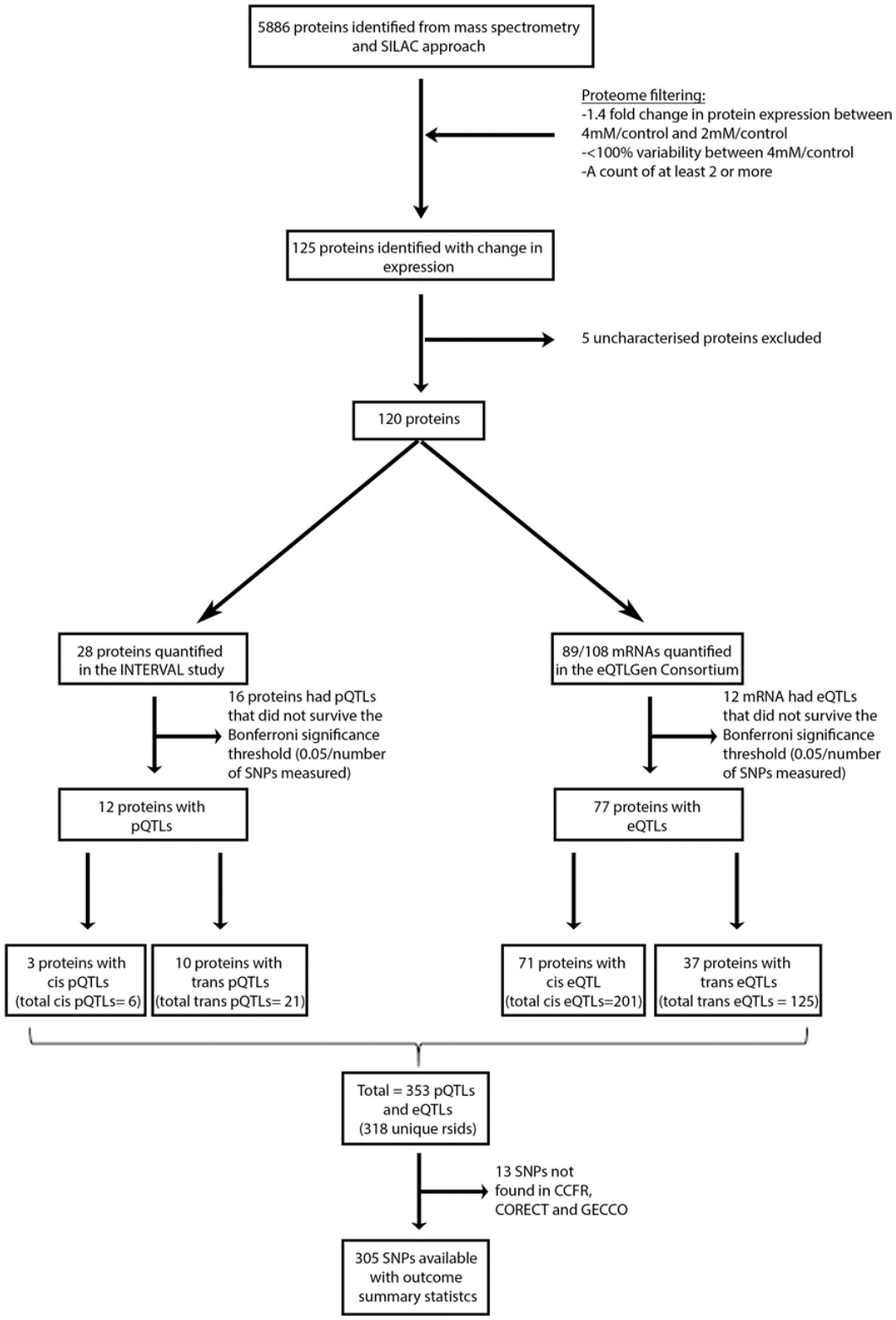
551 $2.50x10^{-4}$ and OR 0.84, 95% CI: 0.75-0.94, P value: $1.17x10^{-3}$, respectively) (Figure 2, Table 1). For
552 *FADS2* and *POLD2*, results were consistent using other MR methods (weighted median, weighted
553 mode and simple mode) and the MR Egger test shows no evidence of pleiotropy (S6 Table,
554 Supplementary Figure 1). From the cis eQTL analysis, only results for *FADS2* survived the Bonferroni
555 significance threshold.

556 Proteins instrumented by trans eQTLs include ribonucleoside-diphosphate reductase subunit M2
557 (*RRM2*), stathmin-1 (*STMN1*) and lipin 1 (*LPIN1*). An increase in *RRM2* was estimated to increase the
558 risk of cancer incidence (OR 3.33, 95% CI: 2.46-4.50, P value: $6.25x10^{-15}$ per SD increase in mRNA
559 expression) whereas an increase in *STMN1* and *LPIN1* was associated with decreases in the risk of
560 CRC incidence (OR 0.72, 95% CI: 0.54-0.97, P value: 0.03 and OR 0.40, 95% CI: 0.32-0.50, P value:
561 $5.50x10^{-16}$ per SD increase in mRNA expression, respectively). From the trans eQTL analysis, results
562 for *RRM2* and *LPIN1* both survived the Bonferroni significance threshold.

563 For proteins instrumented by pQTLs, ADP ribosylation factor interacting protein 2 (ARFIP2) proxied
564 using a trans pQTL conferred an increased risk of CRC incidence (OR 1.15, 95% CI: 1.01-1.29, P value:
565 0.03 per SD increase in protein expression).

566 Overall, the directions of effects between *HLA-A*, *MCM6*, *RRM2* and ARFIP2 and CRC risk obtained
567 from our MR analysis concur with those anticipated given the protective role of aspirin on CRC and
568 the effect of aspirin treatment on expression of these proteins. Aspirin reduces the protein
569 expression of HLA-A, MCM6, RRM2 and ARFIP2 (fold change in protein expression with 4mM aspirin
570 treatment compared to control: 0.55, 0.65, 0.36 and 0.69, respectively, Table 1) and aspirin intake is
571 associated with a decreased risk of CRC (2–4). Our MR analysis shows that increased expression of
572 these proteins is associated with an increased risk of CRC incidence. Taken together, our results
573 indicate that a possible mechanism through which aspirin decreases the risk of CRC incidence is
574 through the downregulation of HLA-A, MCM6, RRM2 and ARFIP2. The direction of effect was less
575 consistent for the other 4 proteins (FADS2, POLD2, STMN1 and LPIN1) showing opposite results to
576 what we would expect based on the proteomic results (Table 1).

577

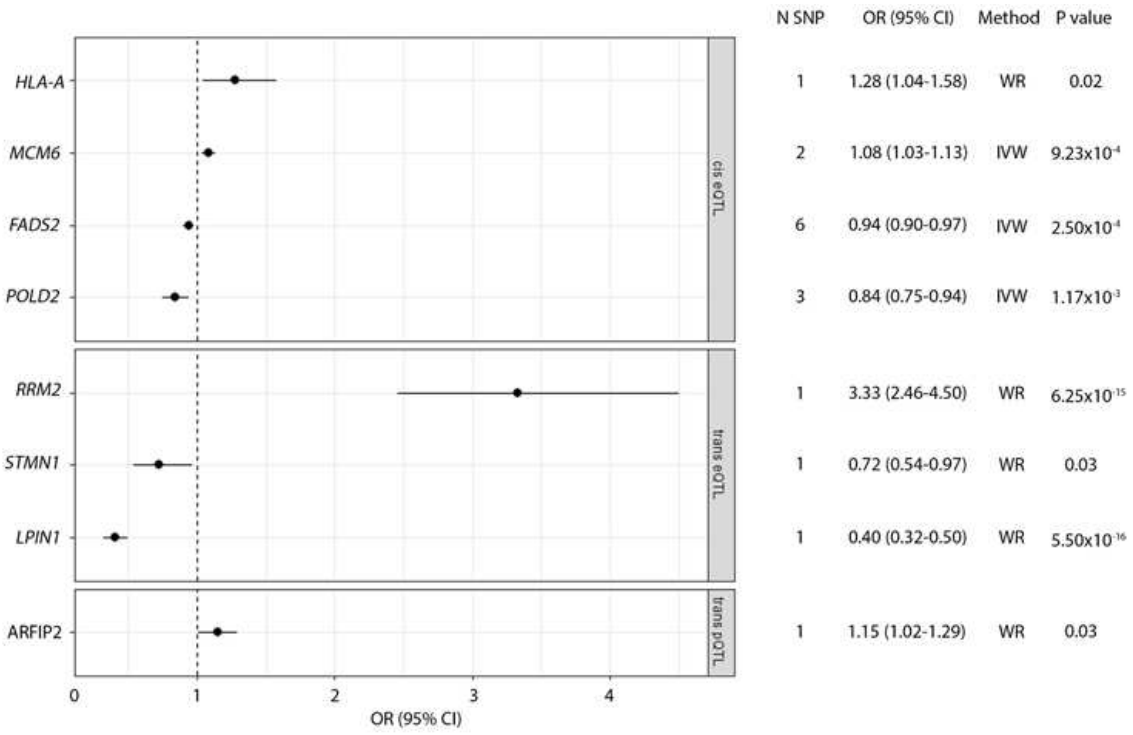


578

579 Figure 1- Flow diagram of SNP selection. 5886 proteins were identified using the SILAC proteomic approach. After applying
580 a threshold, 125 proteins appear to be regulated by aspirin treatment, of which 5 were uncharacterised proteins and were
581 therefore excluded from the analysis. In total, 12 proteins and 77 mRNAs had been quantified and had pQTLs/eQTLs below

582 the Bonferroni significance threshold. Overall, summary statistics for 353 pQTLs and eQTLs were available, of which
583 summary statistics for 305 of the SNPs was also present in the CCFR, CORECT and GECCO consortia.

584



585

586 Figure 2- Forest plot of mRNA/protein associations with CRC incidence at a P value of <0.05. The upper box presents results
587 using cis eQTLs, followed by trans eQTLs and finally trans pQTLs. Each dot on the plot represents the change in OR of CRC
588 incidence per SD increase in mRNA/protein expression and the horizontal lines either side of the dot represent the 95%
589 confidence intervals. The dotted line represents a null association between expression and cancer incidence. The number
590 of SNPs used as instruments as well as the OR, the method and P value of association are also reported. Abbreviations: N
591 SNP, number of SNPs; OR, odds ratio; CI, confidence intervals; IVW, inverse-variance weighted; WR, Wald ratio.

Table 1- MR results of the 8 proteins associated with CRC incidence

					Association of predicted expression with CRC risk					Fold change of protein expression in response to aspirin		
Gene	**Instrument**	**N SNP**	**Variance explained R^2 (%)**	**Method**	**OR**	**LCI**	**UCI**	**P value**	**Effect**	**2mM vs Control**	**4mM vs Control**	**Effect**
FADS2	cis eQTL	6	2.29	IVW	0.94	0.90	0.97	$2.5x10^{-4}$	□	0.61	0.26	□
MCM6	**cis eQTL**	**2**	**3.85**	**IVW**	**1.08**	**1.03**	**1.13**	**$9.23x10^{-4}$**	□	**0.59**	**0.65**	□
POLD2	cis eQTL	3	0.05	IVW	0.84	0.75	0.94	$1.73x10^{-3}$	□	0.54	0.35	□
HLA-A	**cis eQTL**	**1**	**5.95**	**WR**	**1.28**	**1.04**	**1.58**	**0.02**	□	**0.55**	**0.64**	□
LPIN1	trans eQTL	1	0.08	WR	0.40	0.32	0.50	$5.50x10^{-16}$	□	0.65	0.64	□
RRM2	**trans eQTL**	**1**	**0.19**	**WR**	**3.33**	**2.46**	**4.50**	**$6.52x10^{-15}$**	□	**0.33**	**0.36**	□
STMN1	trans eQTL	1	0.04	WR	0.72	0.54	0.97	0.03	□	0.47	0.61	□
ARFIP2	**trans pQTL**	**1**	**0.09**	**WR**	**1.15**	**1.01**	**1.29**	**0.03**	□	**0.67**	**0.69**	□

The table shows the inverse-variance weighted (IVW) or Wald ratio (WR) results for the 7 proteins associated with CRC incidence. The results indicate the change in OR of CRC incidence per unit increase in mRNA or protein expression (z-score or standard deviation, respectively). Results that are consistent with aspirins' effect on protein expression are in bold font. Abbreviations: N SNP, number of SNPs; OR, odds ratio; LCI, lower confidence interval; UCI, upper confidence interval; SE, standard error; IVW, inverse-variance weighted; WR, Wald ratio.

Discussion

Evidence for the use of aspirin in the prevention of CRC is increasing (2–5). However, the mechanism through which it functions is still not fully understood. By combining both a proteomic-based approach as well as an MR analysis, our results provide mechanistic insights into how aspirin could decrease the risk of CRC.

Using a SILAC-based proteomics approach, 120 proteins appear to be regulated at 24 hours by 4mM and 2mM aspirin treatment. Genetic variants (pQTLs and eQTLs) were identified and used to proxy for protein and mRNA expression levels of the identified proteins to test for evidence of a causal effect on CRC incidence. When no pQTL was available for a protein, eQTLs were used instead.

Overall, 4 cis eQTLs, 3 trans eQTLs and 1 trans pQTL were associated with cancer incidence at a P value < 0.05. Increased expression of *HLA-A* and *MCM6* proxied by cis eQTLs were associated with an increase in the risk of CRC incidence and an increase in *RRM2* and ARFIP2 (proxied by a trans eQTL and trans pQTL, respectively) also conferred an increased risk. Therefore, suppressing the expression of these four proteins could decrease the risk of CRC. As the proteomic results showed that aspirin treatment decreases the expression of these proteins, this could be a potential mechanism by which aspirin reduces the risk of CRC. However, only results for *RRM2* survive the Bonferroni significance threshold, indicating that further studies are required to verify these results.

The proteins MCM6 and RRM2 are both involved in repair of DNA damage. MCM6 is part of a helicase complex involved in unwinding DNA and is involved in repair of double stranded breaks (DSBs) in homologous recombination through interaction with RAD51. This interaction is required for chromatin localisation and formation of foci for DNA damage recovery (42). Likewise, RRM2 is part of a protein complex called ribonucleotide reductase which catalyses the biosynthesis of dNTPs and is therefore required for DNA replication and damage repair (43).

Cancer cells commonly lose the DNA damage response, which results in the accumulation of mutations that may be oncogenic (44). Because of this, tumour cells end up relying on a reduced number of repair pathways and are therefore more sensitive to inhibition of DNA damage repair pathways when compared to normal cells which have full capability of DNA repair (45). Drugs that target these other pathways have been shown to selectively kill the cancer cells which is known as synthetic lethality (46,47). It may be that by reducing the expression of DNA repair proteins, which combined with DNA damage response proteins that are already mutated during tumour progression, aspirin can induce cell death in the developing tumour cells reducing the risk of developing cancer.

The MR results for the proteins ARFIP2 and HLA-A also concur with our SILAC proteomic results. ARFIP2 is a protein previously shown to play a role in membrane ruffling and actin polymerization, therefore regulating the actin cytoskeleton (48). The remodelling of the actin cytoskeleton is known to be involved in cancer metastasis (49). This is of particular interest as aspirin reduces the odds of colorectal adenocarcinoma metastasis by 64% (OR:0.36 (95% CI: 0.18-0.74)) (50) and this may be through the reduction in ARFIP2 expression. With regards to HLA-A expression and cancer risk, results from a cohort study showed that aspirin was more chemopreventative in tumours that expressed HLA class I antigen (which includes HLA-A, HLA-B and HLA-C) (rate ratio (RR) 0.53, 95% CI: 0.38-0.74) and this association was no longer apparent in tumours that lacked expression of this protein (15). Our MR analysis showed that an increase in *HLA-A* was associated with increased CRC risk, and that aspirin may reduce this risk through a reduction in HLA-A expression, however further investigation is required before any conclusions can be drawn.

Our MR analysis results also showed that increased mRNA expression of *FADS2*, *POLD2*, *LPIN1* and *STMN1* all decreased the risk of CRC, indicating that decreased expression increases the risk of cancer. Our proteomic results showed that aspirin decreases the expression of these proteins and aspirin is known to decrease cancer risk. The exact meaning behind the inconsistencies in direction of effect is unclear but may be related to the dosage used in this study. A randomized trial of aspirin to prevent adenomas showed that lower doses reduced adenoma risk more than higher doses, suggesting that lower doses of aspirin may affect mRNA/protein expression differently than higher

doses (51,52). Furthermore, the genetic instruments used to proxy for *POLD2*, *LPIN1* and *STMN1* explain little of the variance in mRNA expression (0.05, 0.08 and 0.04%, respectively) indicating that SNPS that explain more of the variance are required before any conclusions can be made.

Further limitations also exist in our analysis. Firstly, the exact correlation between eQTLs and pQTLs has not been fully determined. Secondly, it is difficult to interpret results using trans eQTLs and pQTLs without clear confirmation that these SNPs directly influence the gene/protein expression. It may be that they indirectly influence expression, for example, trans eQTLs may regulate gene expression by affecting expression of a nearby cis gene which is in fact a transcription factor that is regulating the expression of the trans gene (53). Thirdly, both the pQTL and eQTL associations were carried out using blood samples or PBMCs (19,20), therefore these SNPs estimate changes in gene and protein expression in circulating immune cells only. As found by the Genotype-Tissue Expression (GTEx) study, cis eQTLs are either shared across tissues or are specific to a small number of tissues (54). Therefore, the use of these eQTLs and pQTLs measured in the blood may not be fully suitable as proxies for mRNA and protein expression in the epithelium of the colon and rectum. Furthermore, the units for the eQTLs and pQTLs represent SD changes in expression, making interpretation of the results difficult. However, we can interpret the direction of effect as well as the statistical significance of the association (P values) for these analyses. Moreover, pQTLs and eQTLs could not be identified for 20 of the proteins found to be regulated by aspirin in our proteomic approach, therefore we could not test the association of their expression with CRC risk. Finally, apart from the association of *FADS2* with CRC incidence, the other associations proxied by cis eQTLs found by our study are not below the Bonferroni threshold of significance (P value $\leq 4.63 \times 10^{-4}$).

MR is commonly used to proxy for a drug's effect on risk of various outcomes after identification of its target. Genetic variants that predict lower function of 3-hydroxy-3-methylglutaryl coenzyme A (HMG-CoA) reductase are commonly used to investigate the effect of lowering LDL cholesterol via the use of statins on outcomes such as ovarian cancer, Alzheimer's disease or coronary heart disease (55–57). These studies involve investigation of a drug's effect via a known target on an outcome. However, this approach would be difficult to apply in the case of drugs with pleiotropic targets such

675 as aspirin. Therefore, in order to identify all possible targets of aspirin, a proteomic approach was
676 firstly applied and targets that may affect risk of cancer were identified through using MR. To our
677 knowledge, this is the first study that combines basic science and MR to generate hypotheses of a
678 drug's mechanism of action in cancer.

679 Further experiments need to be conducted to confirm the effect of aspirin on gene and protein
680 expression and the consequent effect this may have on hypothesised pathways such as DNA repair
681 before definitive conclusions can be made. However, the potential of this unbiased approach to gain
682 mechanistic insight is clear, allowing hypothesis driven research will better inform the clinical use of
683 aspirin for the prevention of CRC.

684 Acknowledgements:

685 ASTERISK: We are very grateful to Dr. Bruno Buecher without whom this project would not have
686 existed. We also thank all those who agreed to participate in this study, including the patients and
687 the healthy control persons, as well as all the physicians, technicians and students.

688 CLUE: We appreciate the continued efforts of the staff members at the Johns Hopkins George W.
689 Comstock Center for Public Health Research and Prevention in the conduct of the CLUE II study. We
690 thank the participants in CLUE. Cancer incidence data for CLUE were provided by the Maryland
691 Cancer Registry, Center for Cancer Surveillance and Control, Maryland Department of Health, 201 W.
692 Preston Street, Room 400, Baltimore, MD 21201, http://phpa.dhmh.maryland.gov/cancer, 410-767-
693 4055. We acknowledge the State of Maryland, the Maryland Cigarette Restitution Fund, and the
694 National Program of Cancer Registries of the Centers for Disease Control and Prevention for the
695 funds that support the collection and availability of the cancer registry data.

696 COLON and NQplus: the authors would like to thank the COLON and NQplus investigators at
697 Wageningen University & Research and the involved clinicians in the participating hospitals.

698 CORSA: We kindly thank all those who contributed to the screening project Burgenland against CRC.
699 Furthermore, we are grateful to Doris Mejri and Monika Hunjadi for laboratory assistance.

CPS-II: The authors thank the CPS-II participants and Study Management Group for their invaluable contributions to this research. The authors would also like to acknowledge the contribution to this study from central cancer registries supported through the Centers for Disease Control and Prevention National Program of Cancer Registries, and cancer registries supported by the National Cancer Institute Surveillance Epidemiology and End Results program.

Czech Republic CCS: We are thankful to all clinicians in major hospitals in the Czech Republic, without whom the study would not be practicable. We are also sincerely grateful to all patients participating in this study.

DACHS: We thank all participants and cooperating clinicians, and Ute Handte-Daub, Utz Benscheid, Muhabbet Celik and Ursula Eilber for excellent technical assistance.

EDRN: We acknowledge all the following contributors to the development of the resource: University of Pittsburgh School of Medicine, Department of Gastroenterology, Hepatology and Nutrition: Lynda Dzubinski; University of Pittsburgh School of Medicine, Department of Pathology: Michelle Bisceglia; and University of Pittsburgh School of Medicine, Department of Biomedical Informatics.

EPIC: Where authors are identified as personnel of the International Agency for Research on Cancer/World Health Organization, the authors alone are responsible for the views expressed in this article and they do not necessarily represent the decisions, policy or views of the International Agency for Research on Cancer/World Health Organization.

EPICOLON: We are sincerely grateful to all patients participating in this study who were recruited as part of the EPICOLON project. We acknowledge the Spanish National DNA Bank, Biobank of Hospital Clínic–IDIBAPS and Biobanco Vasco for the availability of the samples. The work was carried out (in part) at the Esther Koplowitz Centre, Barcelona.

Harvard cohorts (HPFS, NHS, PHS): The study protocol was approved by the institutional review boards of the Brigham and Women's Hospital and Harvard T.H. Chan School of Public Health, and

725 those of participating registries as required. We would like to thank the participants and staff of the
726 HPFS, NHS and PHS for their valuable contributions as well as the following state cancer registries for
727 their help: AL, AZ, AR, CA, CO, CT, DE, FL, GA, ID, IL, IN, IA, KY, LA, ME, MD, MA, MI, NE, NH, NJ, NY,
728 NC, ND, OH, OK, OR, PA, RI, SC, TN, TX, VA, WA, WY. The authors assume full responsibility for
729 analyses and interpretation of these data.

730 Kentucky: We would like to acknowledge the staff at the Kentucky Cancer Registry.

731 LCCS: We acknowledge the contributions of Jennifer Barrett, Robin Waxman, Gillian Smith and
732 Emma Northwood in conducting this study.

733 NCCCS I & II: We would like to thank the study participants, and the NC Colorectal Cancer Study staff.

734 NSHDS investigators thank the Biobank Research Unit at Umeå University, the Västerbotten
735 Intervention Programme, the Northern Sweden MONICA study and Region Västerbotten for
736 providing data and samples and acknowledge the contribution from Biobank Sweden, supported by
737 the Swedish Research Council (VR 2017-00650).

738 PLCO: The authors thank the PLCO Cancer Screening Trial screening center investigators and the staff
739 from Information Management Services Inc and Westat Inc. Most importantly, we thank the study
740 participants for their contributions that made this study possible.

741 SCCFR: The authors would like to thank the study participants and staff of the Hormones and Colon
742 Cancer and Seattle Cancer Family Registry studies (CORE Studies).

743 SEARCH: We thank the SEARCH team.

744 SELECT: We thank the research and clinical staff at the sites that participated on SELECT study,
745 without whom the trial would not have been successful. We are also grateful to the 35,533
746 dedicated men who participated in SELECT.

747 WHI: The authors thank the WHI investigators and staff for their dedication, and the study
748 participants for making the program possible. A full listing of WHI investigators can be found at:

749 http://www.whi.org/researchers/Documents%20%20Write%20a%20Paper/WHI%20Investigator%20

750 Short%20List.pdf

751 ## References

752 1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global Cancer Statistics 2018:
753 GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries.
754 CA Cancer J Clin. 2018;68(6):394–424.

755 2. Qiao Y, Yang T, Gan Y, Li W, Wang C, Gong Y, et al. Associations between aspirin use and the
756 risk of cancers: a meta-analysis of observational studies. BioMed Cent Cancer. 2018;18(1):1–
757 57.

758 3. Rothwell PM, Wilson M, Elwin CE, Norrving B, Algra A, Warlow CP, et al. Long-term effect of
759 aspirin on colorectal cancer incidence and mortality: 20-year follow-up of five randomised
760 trials. Lancet. 2010;376(9754):1741–50.

761 4. Rothwell PM, Fowkes FGR, Belch JF, Ogawa H, Warlow CP, Meade TW. Effect of daily aspirin
762 on long-term risk of death due to cancer: Analysis of individual patient data from randomised
763 trials. Lancet. 2011;377(9759):31–41.

764 5. Cook NR, Lee I, Zhang SM, Moorthy MV, Buring JE. Alternate-Day, Low-Dose Aspirin and
765 Cancer Risk: Long-Term Observational Follow-up of a Randomized Trial. Ann Intern Med.
766 2013;159(2):77–85.

767 6. Sciulli MG, Filabozzi P, Tacconelli S, Padovano R, Ricciotti E, Capone ML, et al. Platelet
768 activation in patients with colorectal cancer. Prostaglandins Leukot Essent Fat Acids.
769 2005;72(2):79–83.

770 7. Gurpinar E, Grizzle WE, Piazza GA. COX-independent mechanisms of cancer chemoprevention
771 by anti-inflammatory drugs. 2013;3:181.

772 8. Greenhough A, Smartt HJM, Moore AE, Roberts HR, Williams AC, Paraskeva C, et al. The COX-
773 2/PGE2 pathway: key roles in the hallmarks of cancer and adaptation to the tumour
774 microenvironment. Carcinogenesis. 2009;30(3):377–86.

775 9. Alfonso L, Ai G, Spitale RC, Bhat GJ. Molecular targets of aspirin and cancer prevention. Br J
776 Cancer. 2014;111(1):61–7.

777 10. Bak AW, McKnight W, Li P, Soldato P Del, Calignano A, Cirino G, et al. Cyclooxygenase-
778 independent chemoprevention with an aspirin derivative in a rat model of colonic
779 adenocarcinoma. Life Sci. 1998;62(23):PL 367-373.

780 11. Yu H-G, Huang J-A, Yang Y-N, Huang H, Luo H-S, Yu J-P, et al. The effects of acetylsalicylic acid
781 on proliferation, apoptosis, and invasion of cyclooxygenase-2 negative colon cancer cells. Eur
782 J Clin Invest. 2002;32(11):838–46.

783 12. Yin H, Xu H, Zhao Y, Yang W, Cheng J, Zhou Y. Cyclooxygenase-independent effects of aspirin
784 on HT-29 human colon cancer cells, revealed by oligonucleotide microarrays. Biotechnol Lett.
785 2006;28(16):1263–70.

786 13. Borthwick GM, Johnson AS, Partington M, Burn J, Wilson R, Arthur HM. Therapeutic levels of
787 aspirin and salicylate directly inhibit a model of angiogenesis through a Cox- independent
788 mechanism. FASEB J. 2006;20(12):2009–16.

789 14. Domingo E, Church DN, Sieber O, Ramamoorthy R, Yanagisawa Y, Johnstone E, et al.
790 Evaluation of PIK3CA mutation as a predictor of benefit from nonsteroidal anti-inflammatory

drug therapy in colorectal cancer. J Clin Oncol. 2013;31(34):4297–305.

15. Reimers MS, Bastiaannet E, Langley RE, van Eijk R, van Vlierberghe RLP, Lemmens VEP, et al. Expression of HLA Class I Antigen, Aspirin Use, and Survival After a Diagnosis of Colon Cancer. JAMA Intern Med. 2014;174(5):732–9.

16. Drew DA, Cao Y, Chan AT. Aspirin and colorectal cancer: the promise of precision chemoprevention. Nat Rev Cancer. 2016;16:173–86.

17. Lawlor DA, Harbord RM, Sterne JAC, Timpson N, Davey Smith G. Mendelian randomization: Using genes as instruments for making causal inferences in epidemiology. Stat Med. 2008;27(8):1133–63.

18. Davey Smith G, Ebrahim S. What can Mendelian randomisation tell us about modifiable behavioural and environmental exposures? Br Med J. 2005;330(7499):1076–9.

19. Sun BB, Maranville JC, Peters JE, Stacey D, Staley JR, Blackshaw J, et al. Genomic atlas of the human plasma proteome. Nature. 2018;558(7708):73–9.

20. Võsa U, Claringbould A, Westra H-J, Jan Bonder M, Deelen P, Zeng B, et al. Unraveling the polygenic architecture of complex traits using blood eQTL meta-analysis. bioRxiv [Internet]. 2018; Available from: https://www.biorxiv.org/content/10.1101/447367v1

21. Zheng J, Haberland V, Baird D, Walker V, Haycock P, Gutteridge A, et al. Phenome-wide Mendelian randomization mapping the influence of the plasma proteome on complex diseases. bioRxiv [Internet]. 2019; Available from: https://www.biorxiv.org/content/10.1101/627398v1

22. Cole BF, Logan RF, Halabi S, Benamouzig R, Sandler RS, Grainge MJ, et al. Aspirin for the chemoprevention of colorectal adenomas: Meta-analysis of the randomized trials. Vol. 101, Journal of the National Cancer Institute. 2009. p. 256–66.

23. Walther A, Johnstone E, Swanton C, Midgley R, Tomlinson I, Kerr D. Genetic prognostic and predictive markers in colorectal cancer. Nat Rev Cancer. 2009;9(7):489–99.

24. Paraskeva C, Finerty S, Mountford RA, Powell SC. Specific cytogenetic abnormalities in two new human colorectal adenoma-derived epithelial cell lines. Cancer Res. 1989;49(5):1282–6.

25. Browne SJ, Williams AC, Hague A, Butt AJ, Paraskeva C. Loss of APC protein expressed by human colonic epithelial cells and the appearance of a specific low-molecular-weight form is associated with apoptosis in vitro. Int J Cancer. 1994;59(1):56–64.

26. Greenhough A, Wallam CA, Hicks DJ, Moorghen M, Williams AC, Paraskeva C. The proapoptotic BH3-only protein Bim is downregulated in a subset of colorectal cancers and is repressed by antiapoptotic COX-2 / PGE 2 signalling in colorectal adenoma cells. Oncogene. 2010;29(23):3398–410.

27. Baker SJ, Preisinger AC, Jessup JM, Paraskeva C, Markowitz S, Willson JK V, et al. P53 Gene Mutations Occur in Combination With 17P Allelic Deletions As Late Events in Colorectal Tumorigenesis. Cancer Res. 1990;50(23):7717–22.

28. Trinkle-Mulcahy L, Boulon S, Lam YW, Urcia R, Boisvert F-M, Vandermoere F, et al. Identifying specific protein interaction partners using quantitative mass spectrometry and bead proteomes. J Cell Biol. 2008;183(2):223–39.

29. Greenhough A, Bagley C, Heesom KJ, Gurevich DB, Gay D, Bond M, et al. Cancer cell adaptation to hypoxia involves a HIF-GPRC5A-YAP axis. EMBO Mol Med. 2018;10(9):e8699.

30. Yang W, Chung YG, Kim Y, Kim T-K, Keay SK, Zhang C-O, et al. Quantitative proteomics identifies a beta-catenin network as an element of the signaling response to Frizzled-8

protein-related antiproliferative factor. Mol Cell Proteomics. 2011;10(6):M110.007492.

31. Angelantonio E Di, Thompson SG, Kaptoge S, Moore C, Walker M, Armitage J, et al. Efficiency and safety of varying the frequency of whole blood donation (INTERVAL): a randomised trial of 45000 donors. Lancet. 2017;390(10110):2360–71.

32. Sinclair JK, Taylor PJ, Hobbs SJ. Alpha Level Adjustments for Multiple Dependent Variable Analyses and Their Applicability – A Review. Int J Sport Sci Eng. 2013;07(01):17–20.

33. Huyghe JR, Bien SA, Harrison TA, Kang HM, Chen S, Schmit SL, et al. Discovery of common and rare genetic risk variants for colorectal cancer. Nat Genet. 2019;51:76–87.

34. Schumacher FR, Schmit SL, Jiao S, Edlund CK, Wang H, Zhang B, et al. Genome-wide association study of colorectal cancer identifies six new susceptibility loci. Nat Commun. 2015;6:7138.

35. Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, et al. Genome-wide genetic data on ~500,000 UK Biobank participants. bioRxiv [Internet]. 2017; Available from: https://www.biorxiv.org/content/biorxiv/suppl/2017/07/20/166298.DC1/166298-1.pdf

36. Hemani G, Zheng J, Elsworth B, Wade KH, Haberland V, Baird D, et al. The MR-Base platform supports systematic causal inference across the human phenome. Elife. 2018;7:e34408.

37. Burgess S, Butterworth A, Thompson SG. Mendelian Randomization Analysis With Multiple Genetic Variants Using Summarized Data. Genet Epidemiol. 2013;37(7):658–65.

38. Slob EAW, Burgess S. A Comparison Of Robust Mendelian Randomization Methods Using Summary Data. Genet Epidemiol. 2020;44(4):313–29.

39. Hartwig FP, Davey Smith G, Bowden J. Robust inference in summary data Mendelian randomization via the zero modal pleiotropy assumption. Int J Epidemiol. 2017;46(6):1985–98.

40. Bowden J, Davey Smith G, Haycock PC, Burgess S. Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. Genet Epidemiol. 2016;40(4):304–14.

41. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: Effect estimation and bias detection through Egger regression. Int J Epidemiol. 2015;44(2):512–25.

42. Huang J, Luo H-L, Pan H, Qiu C, Hao T-F, Zhu Z-M. Interaction between RAD51 and MCM complex is essential for RAD51 foci forming in colon cancer HCT116 cells. Biochem. 2018;83(1):69–75.

43. Chen CW, Li Y, Hu S, Zhou W, Meng Y, Li Z, et al. DHS (trans−4,4′-dihydroxystilbene) suppresses DNA replication and tumor growth by inhibiting RRM2 (ribonucleotide reductase regulatory subunit M2). Oncogene. 2018;38(13):2364–79.

44. Li XL, Zhou J, Chen ZR, Chng WJ. P53 mutations in colorectal cancer- molecular pathogenesis and pharmacological reactivation. World J Gastroenterol. 2015;21(1):84–93.

45. Brown JS, O'Carrigan B, Jackson SP, Yap TA. Targeting DNA repair in cancer: Beyond PARP inhibitors. Cancer Discov. 2017;7(1):20–37.

46. Hosoya N, Miyagawa K. Targeting DNA damage response in cancer therapy. Cancer Sci. 2014;105(4):370–88.

47. Ashworth A. A synthetic lethal therapeutic approach: Poly(ADP) ribose polymerase inhibitors for the treatment of cancers deficient in DNA double-strand break repair. J Clin Oncol. 2008;26(22):3785–90.

48. D'Souza-Schorey C, L.Boshans R, McDonough M, D.Stahl P, Aelst L Van. A role for POR1, a Rac1-interacting protein, in ARF6-mediated cytoskeletal rearrangements. EMBO J. 1997;16(17):5445–54.

49. Fife CM, McCarroll JA, Kavallaris M. Movers and shakers: cell cytoskeleton in cancer metastasis. Br J Pharmacol. 2014;171(24):5507–23.

50. Rothwell PM, Wilson M, Price JF, Belch JFF, Meade TW, Mehta Z. Effect of daily aspirin on risk of cancer metastasis: A study of incident cancers during randomised controlled trials. Lancet. 2012;379(9826):1591–601.

51. Benamouzig R, Uzzan B, Deyra J, Martin A, Girard B, Little J, et al. Prevention by daily soluble aspirin of colorectal adenoma recurrence: 4-Year results of the APACC randomised trial. Gut. 2012;61(2):255–61.

52. Baron JA, Cole BF, Sandler RS, Haile RW, Ahnen D, Bresalier R, et al. A Randomized Trial of Aspirin to Prevent Colorectal Adenomas. N Engl J Med. 2003;348(10):891–9.

53. Yao C, Joehanes R, Johnson AD, Huan T, Liu C, Freedman JE, et al. Dynamic Role of trans Regulation of Gene Expression in Relation to Complex Traits. Am J Hum Genet. 2017;100(4):571–80.

54. Aguet F, Brown AA, Castel SE, Davis JR, He Y, Jo B, et al. Genetic effects on gene expression across human tissues. Nature. 2017;550(7675):204–13.

55. Yarmolinsky J, Bull CJ, Vincent EE, Robinson J, Walther A, Smith GD, et al. Association Between Genetically Proxied Inhibition of HMG-CoA Reductase and Epithelial Ovarian Cancer. JAMA. 2020;323(7):646–55.

56. Benn M, Nordestgaard BG, Frikke-schmidt R, Tybjærg-Hansen A. Low LDL cholesterol, PCSK9 and HMGCR genetic variation, and risk of Alzheimer's disease and Parkinson's disease: Mendelian randomisation study. BMJ. 2017;357:j3170.

57. Ference BA, Majeed F, Penumetcha R, Flack JM, Brook RD. Effect of Naturally Random Allocation to Lower Low-Density Lipoprotein Cholesterol on the Risk of Coronary Heart DiseaseMediatedbyPolymorphisms in NPC1L1, HMGCR,or Both: A 2 × 2 Factorial Mendelian Randomization Study. J Am Coll Cardiol. 2015;65(15):1552–61.