

1 Combinatorial, additive and dose-dependent drug- 2 microbiome associations

3
4 Sofia K. Forslund^{1,2,3,4,5*}, Rima Chakaroun^{6*}, Maria Zimmermann-Kogadeeva^{1*}, Lajos
5 Markó^{2,4,5*}, Judith Aron-Wisnewsky^{7,8*}, Trine Nielsen^{9*}, Lucas Moitinho-Silva¹, Thomas S. B.
6 Schmidt¹, Gwen Falony¹⁵, Sara Vieira-Silva¹⁵, Solia Adriouch⁷, Renato J. Alves^{1,10}, Karen Ass-
7 mann⁷, Jean-Philippe Bastard^{11,12}, Till Birkner^{2,3}, Robert Caesar¹³, Julien Chilloux¹⁴, Luis
8 Pedro Coelho¹, Leopold Fezeu¹⁶, Nathalie Galleron¹⁷, Gerard Helft¹⁸, Richard Isnard¹⁸, Boyang
9 Ji¹⁹, Michael Kuhn¹, Emmanuelle Le Chatelier¹⁷, Antonis Myridakis¹⁴, Lisa Olsson¹³, Nicolas
10 Pons¹⁷, Edi Prifti^{7,20,21}, Benoit Quinquis¹⁷, Hugo Roume¹⁷, Joe-Elie Salem²², Nataliya
11 Sokolovska⁷, Valentina Tremaroli¹³, Mireia Valles-Colomer¹⁵, Christian Lewinter²³, Nadja B
12 Søndertoft⁹, Helle Krogh Pedersen⁹, Tue H Hansen⁹, The MetaCardis Consortium^{**}, Jens Peter
13 Gøtze²⁴, Lars Køber²³, Henrik Vestergaard^{9,25}, Torben Hansen⁹, Jean-Daniel Zucker^{7,20,21},
14 Serge Hercberg¹⁶, Jean-Michel Oppert⁸, Ivica Letunic^{1,26}, Jens Nielsen¹⁹, Fredrik Bäckhed^{9,13},
15 S. Dusko Ehrlich¹⁷, Marc-Emmanuel Dumas^{14, 27,28,29}, Jeroen Raes¹⁵, Oluf Pedersen⁹, Karine
16 Clément^{7,8+}, Michael Stumvoll^{6,30+}, Peer Bork^{1,3+}

17 * contributed equally

18 + corresponding authors

19 ** MetaCardis Consortium Collaborators

20 *Chloe Amouyal, Ehm Astrid Andersson Galijatovic, Fabrizio Andreelli, Olivier Barthelemy, Jean-Paul Ba-*
21 *tisse, Eugeni Belda, Magalie Berland, Randa Bittar, Hervé Blottière, Frederic Bosquet, Rachid Boubrit,*
22 *Olivier Bourron, Mickael Camus, Dominique Cassuto, Cecile Ciangura, Jean-Philippe Collet, Maria-Car-*
23 *lota Dao, Morad Djebbar, Angélique Doré, Line Engelbrechtsen, Soraya Fellahi, Sebastien Fromentin, Pilar*
24 *Galan, Dominique Gauguier, Philippe Giral, Agnes Hartemann, Bolette Hartmann, Jens Juul Holst, Malene*
25 *Hornbak, Lesley Hoyles, Jean-Sebastien Hulot, Sophie Jaqueminet, Niklas Rye Jørgensen, Hanna Julienne,*
26 *Johanne Justesen, Judith Kammer, Nikolaj Karup, Mathieu Kerneis, Jean Khemis, Ruby Kozlowski, VÉR-*
27 *onique Lejard, Florence Levenez, Lea Lucas-Martini, Robin Massey, Laura Martinez-Gili, Nicolas Maziers,*
28 *Jonathan Medina-Stamminger, Gilles Montalescot, Sandrine Moutel, Ana Luisa Neves, Michael Olanipekun,*
29 *Laetitia Pasero Le Pavin, Christine Poitou, Françoise Pousset, Laurence Pouzoulet, Andrea Rodriguez-*
30 *Martinez, Christine Rouault, Johanne Silvain, Mathilde Svendstrup, Timothy D Swartz, Thierry Vanduyven-*
31 *boden, Camille Vatier, Stefanie Walther.*

32
33

34 **Author Affiliations:**

35 1 Structural and Computational Biology, European Molecular Biology Laboratory, Heidelberg, Ger-
36 many

37 2 Experimental and Clinical Research Center, a cooperation of Charité-Universitätsmedizin and the
38 Max-Delbrück Center, Berlin, Germany

39 3 Max Delbrück Center for Molecular Medicine (MDC), Berlin, Germany

40 4 Charité – Universitätsmedizin Berlin, Berlin, Germany

41 5 Berlin Institute of Health (BIH), Berlin, Germany

42 6 Medical Department III - Endocrinology, Nephrology, Rheumatology, University of Leipzig Medical
43 Center, Leipzig, Germany

44 7 Sorbonne Université, INSERM, Nutrition and obesities; systemic approaches (NutriOmics), Paris,
45 France

46 8 Assistance Publique Hôpitaux de Paris, Pitie-Salpêtrière Hospital, Nutrition department, Paris France

47 9 Novo Nordisk Foundation Center for Basic Metabolic Research, Faculty of Health and Medical Sci-
48 ences, University of Copenhagen, Copenhagen, Denmark

49 10 Collaboration for joint PhD degree between EMBL and Heidelberg University, Faculty of Biosci-
50 ences

51 11 Assistance Publique Hôpitaux de Paris, UF Biomarqueurs Inflammatoires et Métaboliques, Bio-
52 chemistry and Hormonology Department, Tenon Hospital, Paris, France

53 12 Sorbonne Université-INSERM UMR S_938, Centre de Recherche Saint-Antoine, Paris, France

54 13 Wallenberg Laboratory, Department of Molecular and Clinical Medicine and Sahlgrenska Center for
55 Cardiovascular and Metabolic Research, University of Gothenburg, Gothenburg, Sweden

56 14 Division of Systems Medicine, Department of Metabolism, Digestion and Reproduction, Faculty of
57 Medicine, Imperial College London, London, United Kingdom

58 15 Center for Microbiology, VIB, Leuven, Belgium

59 16 Sorbonne Paris Cité Epidemiology and Statistics Research Centre (CRESS), U1153 Inserm, U1125,
60 Inra, Cnam, University of Paris 13, Nutritional Epidemiology Research Team (EREN), 93017, Bobigny,
61 France

62 17 Metagenopolis, INRA, AgroParisTech, Université Paris-Saclay, France

63 18 Assistance Publique Hôpitaux de Paris, Pitié-Salpêtrière Hospital, Cardiology department, Paris

64 France

65 19 Department of Biology and Biological Engineering, Chalmers University of Technology, Gothen-

66 burg, Sweden

67 20 Institute of cardiometabolism and Nutrition, Integromics unit, Paris, France

68 21 Sorbonne Université, IRD, Unité de Modélisation Mathématique et Informatique des Systèmes Com-

69 plexes, UMMISCO, Bondy, France

70 22 Assistance Publique Hôpitaux de Paris, Pitié-Salpêtrière Hospital, Department of Pharmacology,

71 UNICO Cardio-oncology Program, CIC-1421; INSERM, Sorbonne Université, Paris, France.

72 23 Department of Cardiology, Rigshospitalet, University of Copenhagen, Copenhagen, Denmark

73 24 Department of clinical biochemistry, Rigshospitalet, University of Copenhagen, Copenhagen, Den-

74 mark

75 25 Department of Medicine, Bornholms Hospital, Rønne,, Denmark

76 26 Biobyte solutions GmbH, Heidelberg, Germany

77 27 Genomic and Environmental Medicine, National Heart & Lung Institute, Faculty of Medicine, Im-

78 perial College London, London, United Kingdom

79 28 European Genomic Institute for Diabetes, CNRS UMR 8199, INSERM UMR 1283, Institut Pasteur

80 de Lille, Lille University Hospital, University of Lille, Lille, France

81 29 McGill University and Genome Quebec Innovation Centre, Montréal, QC, Canada

82 30 Helmholtz Institute for Metabolic, Obesity and Vascular Research (HI-MAG) of the Helmholtz

83 Zentrum München at the University of Leipzig, Leipzig, Germany

84 **Abstract**

85
86 Upon transition from health to cardiometabolic disease (CMD), patients are heavily medicated, leading
87 to increasingly aberrant gut microbiome and serum metabolome and complicating quests for severity
88 and prognostic biomarkers. Through integrated multi-omics analyses of 2,173 European residents (Met-
89 aCardis cohort), we show that the explanatory power of drugs for variability of both host and gut mi-
90 crobiome features exceeds that of disease. We quantify inferred effects of single and combinatorial
91 medications as well as additive effects, shifting metabolome and microbiome towards a healthier state,
92 such as synergistic reduction of serum atherogenic lipoproteins by statins combined with aspirin, or
93 enrichment of intestinal *Roseburia* by diuretics combined with beta-blockers. Several antibiotics exhibit
94 quantitative relationship between number of courses prescribed during recent five years and progression
95 towards a microbiome state associated with CMD severity. We further report a relationship between
96 cardiometabolic drug dosage, improvement in clinical markers and microbiome composition, support-
97 ing direct drug effects. Taken together, our computational framework and resulting resources allow dis-
98 entangling drug from disease effects on host and microbiome features in heavily medicated subjects.
99 Furthermore, the robust CMD signatures identified with our framework provide new hypotheses for
100 drug-host-microbiome interactions in cardiometabolic disease.

101 **Main text**

102 Identifying and quantifying robust gut microbiota contributions to health and disease requires complex
103 technical and statistical frameworks^{1,2} and remains challenging due to many covariates affecting both
104 microbial composition³⁻⁵ and disease. Among covariates, therapeutic drugs^{4,8-10}, such as broadly pre-
105 scribed proton pump inhibitors (PPI)⁶ and type 2 diabetes (T2D) drug metformin⁷, constitute prime ex-
106 amples. These drugs considerably impact the gut microbiota and modulate inflammation¹¹. Furthermore,
107 direct drug-microbial interactions have been demonstrated *in vitro*⁸. For several drugs in a mostly
108 healthy population, their usage explained more variance in microbiota composition than other covariates
109 tested, albeit with small individual effect sizes¹². However, studies in healthy populations^{12,13} are inad-
110 equate for investigating the secondary impacts of drugs in the context of chronic diseases. To robustly

111 disentangle drug-microbiome associations from host and disease factors, large sample sizes and high
112 resolution of clinical phenotypes over a wide range of disease stages and medication are required for
113 statistical power, while accounting for known variables affecting the gut microbiome. Finally, biological
114 effects of drugs are often dose-dependent, yet dose relations have rarely been considered in microbiome
115 studies.

116 To overcome these limitations, we propose a general framework for separating disease from treatment
117 associations in multi-omics cross-sectional studies and apply it to gut metagenomic, host clinical and
118 metabolomic measurements of 2,173 European residents from the MetaCardis cohort (Methods, Ex-
119 tended Data Figure 1, Supplementary Table 1). MetaCardis constitutes a multi-centre (Denmark, France,
120 and Germany) cross-sectional study, with participants ranging from healthy over metabolic syndrome
121 (MetS), severe and morbid obesity, type 2 diabetes (T2D), to those with severe cardiometabolic disease
122 (CMD), e.g. acute and chronic coronary artery disease (CAD), and heart failure (HF), both CAD-asso-
123 ciated and not. Considering current CMD- and other frequently prescribed medications, we investigated
124 drug-host-microbiome associations for eight major therapeutic indications (antidiabetic, antihyperten-
125 sive, antidyslipidemic, antithrombotic, antiarrhythmic agents, gout medication, drugs treating acid-re-
126 flux-related disorders such as PPIs, and antibiotics spanning over 49 individual drug classes (Supple-
127 mentary Tables 2-4)). We further investigated known CMD risk factors (age, sex, body mass index
128 (BMI), diet, smoking), while controlling for variability traceable to the study centres. The most com-
129 monly prescribed CMD drugs were statins (n = 772, 35.5%), beta-blockers (n = 656, 30.2%), metformin
130 (n = 607, 27.9%), aspirin (n = 532, 24.5%), angiotensin converting enzyme (ACE) inhibitors (n = 470,
131 21.6%) and angiotensin II receptor blockers (ARB) (n = 470, 21.6%) reflecting European standards of
132 care in CMD (Supplementary Table 3). Several drugs were taken in combination (Supplementary Table
133 3). We therefore studied individual drug effects, as well as their synergistic and additive interactions in
134 the context of available phenotypic, dietary, and demographic variables, molecular readouts including
135 serum concentrations of lipoproteins, cytokines and metabolites, and taxonomic and functional profiles
136 of the gut microbiome.

137 To quantify the overall impact of medications, we performed multivariate regression of explained vari-
138 ance of host and microbiome data onto total influence of medications, clinical and environmental risk

139 factors and disease status (Methods). All drugs together explain more variation in the microbiome com-
140 position than patient disease group does, or any other factor considered under a conservative estimate
141 (Figure 1a). However, in line with previously reported high individual variability¹⁴, only 1.7 - 9% of
142 variation between subjects is explainable by the factors included in the model, of which 1 - 2.5% are
143 attributable to drug intake, which is comparable to disease status, diet and smoking combined (Figure
144 1a, Supplementary Table 5).

145 To quantify individual drug effects, we implemented a univariate statistical approach to disentangle
146 drugs from disease associations with the gut microbiome and host features. We marked each association
147 fully reducible to one or more non-disease covariates as confounded, considering all frequently pre-
148 scribed CMD drugs, singly and in combination. Thus, features distinguishing patient groups from
149 healthy controls are divided into i) confidently deconfounded features of CMD, ii) ambiguously decon-
150 founded (where both treatment and disease strongly correlate), and iii) confounded (unambiguous drug
151 associations) (Methods, Extended Data Figure 1). A major fraction of naïve associations (e.g. 45% for
152 T2D) between drugs and microbiome or metabolome is attributable to drug intake (Figure 1b, Supple-
153 mentary Table 5). Nonetheless, we recover previously described metabolic disease signatures in micro-
154 biome and metabolome and show these cannot be reduced to treatment effects (Extended Data Figure
155 2, Supplementary Results section 2.3). We thus conclude that, at least for CMD, a drug-conscious ap-
156 proach uncovers true disease associations and is crucial to circumvent highly inflated treatment-con-
157 founded false positives in biomarker discovery.

158 Having quantified the impact of individual drugs, we then disentangled potential direct effects of the
159 medication (where treatment association direction opposes the disease association) from potential se-
160 verity markers (concordant direction of the treatment and disease association). Of 28 cardiometabolic
161 drugs taken by sufficiently many study participants (at least 10 individuals within at least one patient
162 group), the strongest effects on serum metabolome were found for antidiabetic drugs, statins¹¹, beta-
163 blockers, antithrombotic drugs and aspirin (Figure 1c). While drugs with the same indication (i.e. anti-
164 diabetic, antihypertensive) had concordant associations with host features, the impact on the gut micro-
165 biome was more diverse in effect size and direction between these drugs (Figure 1c, Supplementary
166 Tables 6, 7). Our approach recaptured previously reported findings on the impact of antibiotics¹⁵, PPIs,

167 statins¹¹, beta-blockers^{16,17} and metformin (Extended Data Figure 3, Supplementary Table 6,
168 Supplementary Results section 2.3). More importantly, we herein identified novel associations for these
169 reported as well as for other highly prevalent drugs (Supplementary Results section 2.4). For example,
170 we identified aspirin-associated changes in microbial species abundances, as well as shifts in serum
171 lipidome and metabolome associated with improved cardiometabolic health (e.g., depletion of *Rumino-*
172 *coccus gnavus*, *Clostridium glycyrrhizinilyticum* and *Parvimonas micra*, reduction of plasma concen-
173 trations of inflammatory markers (CRP and IL6), decreased levels of pyruvate, glutamate and succinate
174 at comparable significance to that of the aspirin levels detected in serum of medicated subjects; Figure
175 1d2b, Supplementary Table 6, Supplementary Results section 2.4). In addition, γ -butyrobetaine, a re-
176 cently identified proatherogenic intermediate of microbial metabolism¹⁹, is lower in subjects taking as-
177 pirin, revealing a potential complex antiatherogenic effect of the drug beyond its known platelet-inhib-
178 itory functions²⁰. For the known gut modulator metformin, we deduce novel antidiabetic effects possibly
179 related to lowered glutamate levels²¹ ($d = -0.17$, FDR = 0.02), due to reduced microbial glutamate
180 transport ($d = -0.2$, FDR = 0.006). Furthermore, we observe increased microbial vitamin B12 uptake (d
181 = 0.32, FDR=3.65e-6), potentially leading to vitamin B12 deficiency in the host, a known metformin
182 side effect (Supplementary Results section 2.4, Supplementary Table 6). PPIs had the most associations
183 with gut microbiome features (Figure 1c, Supplementary Table 7) including higher prevalence of pre-
184 sumably oral bacteria (Supplementary Table 6), supporting the hypothesized PPI-caused transfer of oral
185 bacteria into the gut upon decreased stomach acidity¹⁷. Single nucleotide variation (SNV) analysis based
186 on large reference cohorts (Supplementary Results section 2.4) revealed increased abundance of usually
187 oral-based strains of *Rothia*, *Haemophilus* and *Streptococcus* species in the gut of subjects taking PPIs,
188 implying that the patient's own oral strains colonize the intestine as gastric acidity weakens²² (Figure
189 1e).

190 Beyond single drugs, the MetaCardis study population enables analysis of combinatorial (polyphar-
191 macy) effects, since 1,300 individuals were prescribed more than one drug (average daily intake of 3
192 drugs with some receiving up to 13 distinct drugs per day) (Figure 2a, Supplementary Table 2). Most
193 common drug combinations include aspirin and statins (437 subjects, 20.1%), beta-blockers and statins
194 (413 subjects, 19%), beta-blockers and aspirin (337 subjects, 15.5%), and the triad of beta-blockers,

195 aspirin and statins (298 subjects, 13.7%), the cornerstone treatment in CAD (Figure 2b, Supplementary
196 Table 3). Polypharmacy in CMD mostly reflects concurrence of metabolic diseases, risk factors, or
197 treatments preventing the recurrence of an atherosclerotic event, but also includes medications co-pre-
198 scribed to reduce side effects, such as PPIs with aspirin and clopidogrel to prevent gastric ulcers and
199 bleeding. Multi-medicated patients often exhibit a more pronounced improvement in disease markers
200 than those receiving either drug alone, consistent with synergistic interactions between drugs (Supple-
201 mentary Table 8). In the T2D group, the most pronounced synergistic effects on the microbiome features
202 were observed for loop diuretics, especially in combination with aspirin, ACE-inhibitors and beta-block-
203 ers, whereas the most pronounced synergistic effects on host features were observed for statins (Figure
204 2c). For example, (i) loop diuretics combined with aspirin, ACE-inhibitors or beta-blockers more
205 strongly enrich microbiome-related health markers²³ including *Roseburia* abundance (combination: $d =$
206 0.46 , $d = 0.51$, $d = 0.36$, correspondingly, single drugs: diuretics $d = 0.27$), (ii) calcium channel blockers
207 taken with statins are associated with lower serum concentrations of atherogenic very low-density lipo-
208 proteins (vLDL) (combination: average $d = -0.17$, single drugs: statin average $d = -0.14$) (Figure 2d).
209 (iii) Taken with metformin or aspirin, statins are associated with lower low, intermediate, and very low-
210 density lipoproteins levels in serum and total body fat mass, while increasing microbiome richness and
211 abundance of Firmicutes and methanogenic bacteria otherwise depleted in the T2D group (Figure 2d,
212 Supplementary Tables 8, 9). These shifts in the microbiome might mediate some of the synergistic drug
213 effects on the host (Supplementary Results section 2.5, Figure 2e, Supplementary Table 10).

214 Next, we investigated additive drug associations. The strongest of those we observed for antibiotics
215 using five-year retrospective exposure (total number of courses). Antibiotics are not used to treat CMD,
216 yet are frequently prescribed due to an increased prevalence of infections in this disease population²⁴.
217 Yet, epidemiological studies link antibiotics with an increased risk for obesity, T2D, metabolic and
218 inflammatory diseases²⁵. We observed that previous antibiotic exposure is significantly (i) associated
219 with lower gut gene richness within the same subject groups (Figure 3a, Spearman $\rho = -0.25$, $P = 3.7e-$
220 5) and, (ii) correlated with total abundance of antimicrobial resistance genes (AMR) in the gut (controls:
221 Spearman $\rho = 0.30$, $P = 9e-7$; T2D subjects: Spearman $\rho = 0.20$, $P = 2e-5$) (Figure 3b). These find-
222 ings imply cumulative, additive shifts upon repeated antibiotic exposure towards a more resistant but

223 less diverse microbiota, which is a hallmark of microbiome signature in obesity, insulin resistance and
224 low-grade inflammation²⁶. The same properties distinguish antibiotics-naïve CMD patients from healthy
225 controls confirming a genuine impact of repeated antibiotic exposures (antibiotics-naïve healthy vs T2D
226 richness $P = 2e-16$; AMR gene abundance $P = 2e-2$). Using principal component analysis (PCA, Sup-
227 plementary Table 11), we show that the first PC of microbiome composition, explaining 45% of varia-
228 tion and correlating with gene richness, is associated both with an additive effect of antibiotics and
229 metabolic impairment following antibiotics exposure (antibiotic effect: controls: Spearman rho = 0.27,
230 $P = 1.7e-5$; T2D subjects: Spearman rho = 0.16, $P = 7e-4$; antibiotics-naïve vs antibiotics treated healthy
231 ($P = 1e-3$) and T2D subjects ($P = 1e-3$)) (Figure 3c). This suggests a link between changes in microbiome
232 richness and structure and the epidemiological findings described above. Multivariate breakdown of
233 these shifts reveals reduced abundance of *Prevotella copri* and *Faecalibacterium prausnitzii*, and an
234 increase in *Bacteroides vulgatus* and *Bacteroides dorei*, abundant genera constituting hallmarks of en-
235 terotypes^{27,28}. Further, while controlling for disease and medication intake, we show that shifts in gut
236 microbial metabolic functions link additive effects of specific antibiotics groups to CMD susceptibility
237 (Supplementary Results section 2.6, Extended Data Figures 4-6, Supplementary Table 12).

238 Alongside recurrent drug exposure, the detailed medication tracking in MetaCardis allows to investigate
239 the effect of dosage on the host and microbiota. For the 20 drugs with sufficient dosage information, we
240 distinguished between dosage-confirmed effects, i.e., features significantly associated both with drug
241 intake (yes/no) and with its dosage; and dosage-unique effects, where dosage analysis revealed associ-
242 ations not captured by other analyses. The drugs with the most features confirmed by dosage analysis
243 were metformin, sulfonylurea, insulin, PPI, gout medications, and statins, whereas the most dosage-
244 unique associations were reported for metformin and statins (Figure 3d, Supplementary Table 13). Thus,
245 statin dosage was more strongly negatively associated with atherogenic vLDL levels in serum, high-
246 lighting the intended dose-dependent lipid lowering effects of this drug class, but also revealed a strong
247 positive association with health-promoting *Roseburia* species in the gut¹¹. Metformin dosage was neg-
248 atively associated with cytokine levels (SDF1 and MIF)^{29,30}, consistent with previous reports of its anti-
249 inflammatory effects. Furthermore, metformin dosage was negatively associated with many Firmicutes
250 and positively with *Bacteroides* (Supplementary Table 13), reflecting a shift between Bact1 and Bact2

251 enterotypes in patients taking higher dosages of metformin, which was also associated with disease,
252 proposing Bact2 enterotype as a severity marker in T2D¹¹ (Figure 3e, f, Supplementary Table 14). For
253 statins, dosage analysis strengthens the reported observation of statins shifting the microbiome towards
254 a healthier state away from Bact2 enterotype¹¹. Moreover, dosage analysis uniquely identified Bact2 and
255 Prev enterotypes as severity markers for beta-blocker usage in individuals with severe and morbid obe-
256 sity (Figure 3e, f, Supplementary Table 14).

257 With stringent analytical approaches, we show that not only medication intake, but also dosage, drug
258 combinations and previous exposure to antibiotics should be captured in human studies to disentangle
259 the drug-host-microbiome interactions in complex diseases. For several drugs, our results identify mi-
260 crobiome shifts associated with medication intake, which might mediate the improvement in clinical
261 markers. Since the nature of our study allows to identify associative and not necessarily causative ef-
262 fects, experimental validation using established animal models (e.g. multimodal effect of low-dose as-
263 pirin or synergistic effects of statin and aspirin or metformin in high-fat fed LDL-receptor-deficient
264 mice) is required to confirm these findings, since controlled clinical trials can be challenging in a pop-
265 ulation with multimorbidity. Disentangling medication effects on the gut microbiome and serum metab-
266 olome, as illustrated here, is the first step towards understanding the systemic effects of drugs at the
267 molecular level. To improve treatment in the context of genetic and microbiome variability, drug-aware
268 molecular markers need to be identified along the transition from health to chronic diseases. Subse-
269 quently, the gut modulation potential of drugs could be harnessed to reverse this progression in a per-
270 sonalized manner.

271 Figure legends

272 **Figure 1. General and specific associations between CMD drugs, host and microbiome.**

273 **a.** Stacked bar charts show variance explained (R squared) by variable group and feature type.

274 **b.** Violin plot representing confounder analysis of features differentially abundant between T2D and
275 control subjects; density along vertical axis represents distribution of effect size, total features per cate-
276 gory listed. “Naïve associations” (yellow, two-sided MWU FDR < 0.1) are either confounded or am-
277 biguously/confidently deconfounded (blue, purple and red violins; post-hoc test for covariates). Green
278 violins show breakdown of significant drug confounders by drug category.

279 **c.** Hierarchical clustering of host (top) and microbiome (bottom) features associated with each drug in
280 at least one patient group. Features separate into potential drug effects (discordant with disease associ-
281 ations) and severity markers (concordant with disease associations).

282 **d.** Scatterplot (top) shows effect sizes (Cliff's delta) of confidently deconfounded associations between
283 aspirin usage and serum metabolome, host phenotype and microbiome features, versus effect size of
284 disease when comparing patients and healthy controls within each clinical group. A subset of features
285 is highlighted for interpretation (bottom).

286 **e.** Cuneiform plot shows change in abundance of bacterial species in the gut in subjects taking/not taking
287 PPIs (controlling for other drugs and demographic factors) in each clinical group separately, and for all
288 subjects pooled together. Rows marked “SNV” show whether oral strain single nucleotide markers are
289 significantly (two-sided MWU FDR < 0.1) enriched over gut strain markers in subjects taking PPIs,
290 controlling for abundance of each species. Marker direction, color and size denote the sign and value of
291 Cliff's delta standardized effect size; opaque markers are significantly altered (two-sided MWU FDR <
292 0.1; passing all confounder checks). Bacteria are shown if their abundance is significantly altered under
293 PPI consumption, and there are SNPs distinguishing oral from gut strains in HMP samples. (See Sup-
294 plementary Tables 5-7).

295 **Figure 2. Combinatorial impacts of CMD drugs.**

296 **a.** Number of CMD patients receiving each drug (horizontal axis) singly or in combination with a spec-
297 ified number (stacked bars) of other drugs.

298 **b.** The thirty most common drug combinations represented as a graph. Node size is proportional to the
299 number of combinations per drug; drug pairs are represented by solid lines; drug triplets are represented
300 by distinct dotted/dashed lines. Edge width is proportional to the number of users per combination; edge
301 colour corresponds to number of significant drug associations.

302 **c.** Heatmap shows number of features (separated into host (bottom, green) and microbiome (top, brown))
303 affected by each drug combination more strongly than by single drugs among T2D patients. Diagonal
304 values show number of features affected by each drug alone among T2D patients. Shown are associa-
305 tions that were deconfounded, discordant with the disease effect and significant (two-sided MWU FDR
306 < 0.1).

307 **d.** Effect size (Cliff's delta) and direction of disease associations (T2D, red), drug combinations (black)
308 and single drugs (other colours) among T2D patients for the combination of statin and metformin, aspi-
309 rin or calcium antagonist. Each line on the horizontal axis corresponds to one association between a
310 feature and a drug combination.

311 **e.** Drug-feature graph showing potential mediation between host and microbiome features. Solid lines
312 represent drug effects on the feature, colour represents direction of the effect. Dashed lines between
313 features indicate potential mediation (general mediation model one-sided $P < 0.1$), colour represents the
314 sign of Pearson's correlation coefficient ($P < 0.1$). (See Supplementary Tables 8, 10).

315
316 **Figure 3. Additive and dose-dependent drug associations with host and microbiome.**

317 Scatterplots show microbiome features (**a.** Gene richness; **b.** Total abundance of antibiotic resistance
318 genes; **c.** The first principal component of gut species composition) significantly associated with the
319 number of antibiotics courses in the last 5 years in control and T2D subjects separately (with lines and
320 gray area representing 95% CI for linear regression). Boxplots (box showing median and quartiles,
321 whiskers 1.5 interquartile range, dots outliers) show the comparisons in antibiotics-naïve and antibiotics-
322 exposed controls and T2D subjects, respectively, with pairwise significances (two-sided MWU tests,
323 FDR-adjusted).

324 **d.** Heatmaps show host and microbiome features confirmed by dosage analysis (replicable in a post-hoc
325 test at Spearman $P < 0.05$ excluding wholly unmedicated subjects) (left), or which can be demonstrated
326 only when considering dosage of the medication (right). Features are separated by potential drug effects
327 (discordant with the disease effect) or severity markers (concordant with the disease effect).

328 **e.** Scatterplot shows relationship between drug intake (taking/not taking) effect size (Cliff's delta) and
329 drug dosage (continuous) effect size (Spearman's rho) on enterotype distribution within each patient
330 group. Features significantly affected in either analysis (two-sided, MWU FDR < 0.1) are shown in
331 green (potential drug effects) or purple (potential severity markers). Black circles and text highlight
332 enterotype-drug-patient group associations that are depicted in panel f. Bact1, 2, Bacteroides 1, 2, Prev,
333 Prevotella, Rum, Ruminococcus.

334 **f.** Coloured areas represent the stacked enterotype prevalence along the drug dosage axis, with lines
335 calculated as a fraction of enterotypes in patient subgroups for which drug dosage fall within the corre-
336 sponding value range. Each dot represents a patient taking specific drug dose and classified into one of
337 the four enterotypes. Random noise was added to the dot coordinates for better visualization. (See Sup-
338 plementary Tables 11-14).

339 Extended Data Figures

340 **ED Figure 1. A post-hoc testing approach for deconfounding univariate biomarker analysis for**
341 **multiple medications and risk factors.** The schematic highlights our covariate control approach. All
342 significant associations between putative drivers (e.g., disease D) and covariates ($C_1 \dots C_n$) to each meas-
343 ured feature ($Y_1 \dots Y_m$) are taken. The outcome of the test is denoted with a_i for a positive outcome
344 (“yes”) and \bar{a}_i for a negative outcome (“no”). A significant predictor is called “confounded” and is fil-
345 tered out in a post-hoc test if there is at least one covariate (e.g., drug treatment or combination) such
346 that the predictor does not add significant predictive capacity beyond the covariate (“confounded”). If
347 no such covariate itself passes the same test (i.e., covariates cannot in turn be shown to have predictive
348 capacity beyond tested predictor), the predictor is considered ambiguous (“ambiguously decon-

349 founded”). Otherwise, the predictor is considered “confidently deconfounded” (we note that “confi-
350 dently deconfounded” is defined as no confounders were found among all covariates measured in our
351 study).

352

353 **ED Figure 2. Previously reported metabolic disease associations are replicated in the MetaCardis**
354 **cohort under drug deconfounding, highlighting systemic inflammation, short-chain fatty acid and**
355 **branched-chain amino acid mechanisms underlying insulin resistance.** Cuneiform plot marker hues
356 and direction show sign of effect size (Cliff’s delta), intensity and size show amplitude of effect size,
357 comparing metabolic diseased proband subsets (horizontal axis) with healthy control subject in the Met-
358 aCardis population for different microbiome, metabolome and host features (vertical axis). Bold and
359 opaque markers show significant associations (two-sided MWU FDR < 0.1) not reducible to any signif-
360 icant drug or demographic confounder. Full associations are found in Supplementary Table 9; here a
361 preselected subset is displayed reflecting previously reported risk and protective factors, validated in
362 MetaCardis. ¹H NMR features are shown with retention time in parentheses, functional modules with
363 GMM or KEGG identifier in parenthesis, analogous for metagenomic species and mOTUs.

364

365 **ED Figure 3. Previously reported drug-microbiome associations are replicated in the MetaCardis**
366 **cohort for metformin and PPI.** Bar plots show the magnitude and direction of effect size (Cliff’s delta)
367 of metformin treatment (left) and PPI treatment (right) on microbiome features. These effects are com-
368 pared to the previously published data from two independent patient cohorts. Only features with direct
369 match on the taxonomic level were included in the comparison¹⁰. Full list of associations is provided in
370 Supplementary Table 6.

371

372 **ED Figure 4. Breakdown of antibiotics association into individual features, selected features**
373 **shown.** Left cuneiform plot (markers show Spearman correlation direction by shape and color, scope
374 by size and color, significance (two-sided MWU FDR < 0.1, deconfounded for other drug and demo-
375 graphic features) by edge opacity) shows association between each feature and total number of antibi-
376 otics courses in CMD groups as well as in healthy controls. Right cuneiform shows whether the same

377 features are significantly different (two-sided MWU FDR < 0.1) between healthy controls and CMD
378 subjects following drug deconfounding (markers show Cliff's delta effect size), requiring significant
379 and deconfounded correlation with number of antibiotic courses demonstrable in at least one proband
380 group and at least one group showing significant and deconfounded alteration compared to healthy con-
381 trols. Core features include increased carriage of possible disease-associated *Ruminococcus gnavus* and
382 various *Clostridia* species, alongside decreased carriage of commensals such as *Faecalibacterium* spe-
383 cies. Full list of associations is provided in Supplementary Table 12.

384

385 **ED Figure 5. Taxonomic changes are validated in a recent intervention cohort.** For bacterial species
386 where an effect on abundance of total antibiotics courses in MetaCardis could be demonstrated (signif-
387 icant at Spearman FDR < 0.1 and deconfounded), where effect of antibiotic intervention could also be
388 tested in a recent antibiotic intervention study³¹, effect sizes are shown here (MetaCardis correlation on
389 vertical axis, intervention log-transformed fold change on horizontal axis). Separate markers are shown
390 for each MetaCardis patient group within which antibiotic effect can be demonstrated. Bold markers
391 achieve significance (FDR < 0.1) in the intervention study as well. For the majority of taxa overlapping
392 between studies, direction of changes matches, consistent with a causal impact of antibiotics on the
393 microbiota in MetaCardis.

394

395 **ED Figure 6. Enterotype likelihood is altered by antibiotics.** Cuneiform shows normalized regression
396 coefficients of logistic models for each 4-class enterotype as a function of antibiotics courses in last 5
397 years, separately for controls and metabolic disease patient groups. All significant (two-sided Wald FDR
398 < 0.1) models show depletion of *Ruminococcus* and *Prevotella* enterotypes, and enrichment for Bac-
399 teroidetes enterotypes; in the case of metabolic disease patients, this is strongest for the low cell count
400 Bacteroidetes 2 enterotype.

401

402 **ED Figure 7. Illustration of flow cytometry gating strategy.** A fixed gating/staining approach was ap-
403 plied³². Both blank and sample solutions were stained with SYBR Green I.

404 **a.** FL1-A/FL3-A acquisition plot of a blank sample (0.85% w/v physiological solution) with gate bound-
405 aries indicated. A threshold value of 2000 was applied on the FL1 channel.

406 **b.** Secondary gating was performed on the FSC-A/SSC-A channels to further discriminate between de-
407 bris/background and microbial events.

408 **c, d./** FL1-A/FL3-A count acquisition of a faecal sample with secondary gating on FSC-A/SSC-A chan-
409 nels based on blank analyses. Total counts were defined as events registered in the FL1-A/FL3-A gating
410 area excluding debris/background events observed in the FSC-A/SSC-A R1 gate. The flow rate was set
411 at 14 microliters per minute and the acquisition rate did not exceed 10,000 events per second. Each panel
412 reflects the events registered during a 30 seconds acquisition period. Cell counts were determined in
413 duplicate starting from a single biological sample.

414

415 **Acknowledgements**

416 We thank the MetaCardis subjects for their participation in the study, and particularly the patient asso-
417 ciations (Alliance du Coeur and CNAO) for their input and interface. We further thank Dr Dominique
418 Bonnefont-Rousselot (Department of Metabolic Biochemistry, Pitié-Salpêtrière hospital) for the analy-
419 sis of plasma lipid profiles. We thank the nurses, technicians, clinical research assistants and data man-
420 agers from the Clinical Investigation Platform at the Institute of Cardiometabolism and Nutrition for
421 patient investigations, the Clinical Investigation Center (CIC) from Pitié-Salpêtrière Hospital and Hu-
422 man Research Center on Nutrition (CRNH Ile de France) as well as the university hospital of Leipzig
423 for investigation of healthy controls. Quanta Medical provided regulatory oversight of the clinical study
424 and contributed to the processing and management of electronic data.

425

426 **Availability of data and materials**

427 Raw shotgun sequencing data that support the findings of this study have been deposited in The Euro-
428 pean Nucleotide Archive (ENA) with accession codes [PRJEB41311, PRJEB38742 and PRJEB37249]
429 with public access. The metadata for all samples are provided in Supplementary Tables 2 and 3. The
430 metadata, processed microbiome and metabolome data and code resource are available under

431 <https://doi.org/10.5281/zenodo.4674360> for download. The source data for the figures and correspond-
432 ing code are provided under <https://doi.org/10.5281/zenodo.4728981>. Generally, access to the MetaCardis
433 data and biosamples is available on a project-by-project basis, where researchers may submit a specific
434 reuse request to the consortium via its coordinator Prof. Karine Clément and thereafter formally be
435 granted access; thus, satisfying requirements both of informed consent and of open science. For further
436 data-related questions, contact P.B. For clinical cohort-related questions, contact K.C.

437

438 **Code availability**

439 The novel drug-aware univariate biomarker testing pipeline is available as an R package (metadecon-
440 foundR; Birkner et al., manuscript in preparation) on Github ([https://github.com/TillBirkner/meta-](https://github.com/TillBirkner/meta-deconfoundR)
441 [deconfoundR](https://github.com/TillBirkner/meta-deconfoundR)) and under <https://doi.org/10.5281/zenodo.4721078>. The latest version (0.1.8) of
442 this package was used to generate the data shown in this publication. The code used for multivariate
443 analysis based on the VpThemAll package is available under <https://doi.org/10.5281/zenodo.4719526>.
444 The metadata, processed microbiome and metabolome data and code resource are available under
445 <https://doi.org/10.5281/zenodo.4674360> for download. The source data for the figures and correspond-
446 ing code are provided under <https://doi.org/10.5281/zenodo.4728981>.

447 **Sources of funding**

448 This work was supported by European Union's Seventh Framework Program for research, technological
449 development and demonstration under grant agreement HEALTH-F4-2012-305312 (METACARDIS).
450 Part of this work was also supported by the Metagenopolis grant ANR-11-DPBS-0001. Assistance
451 Publique-Hôpitaux de Paris (AP-HP) is the promoter of the clinical investigation (MetaCardis). MED
452 is supported by the NIHR Imperial Biomedical Research Centre.

453 **Author contributions**

454 KC (coordinator), PB, MS, OP, SDE, JR, M-ED, FB and JN conceived the overall objectives and study
455 design of the MetaCardis initiative. SKF, PB developed the present project concept and protocol and

456 supervised the project. MetaCardis cohort recruitment, phenotyping and lifestyle recording were con-
457 ducted by: RC, JA-W, TN, CL, LK, TH, THH, HV, KA and supervised by MS, KC and OP. Data cura-
458 tion was undertaken by: SKF, RC, LM, KA, JA-W, TN. Faecal microbial DNA extraction and shotgun
459 sequencing: NP, ELC, SF. Bacterial cell count measurement: GF, SVS. Serum and urine metabolome
460 profiling: LH, JC, AM, MO. Bioinformatics and statistical analyses: TB, MZ-K, SKF, LS, TSBS, LPC,
461 NS, JZ, EP, SF, RC, SV, GF and BJ. The manuscript was drafted by SKF, RC, MZ-K, LM. All authors
462 participated in the project development, revision of article and approved the final version for publication.

463 **Competing interests**

464 FB is shareholder in Implexion pharma AB. KC is a consultant for Danone Research, LNC therapeutics
465 and CONFO therapeutics for work unassociated with the present study. KC has held a collaborative
466 research contract with Danone Research in the context of MetaCardis project. MB received lecture
467 and/or consultancy fees from AstraZeneca, Boehringer-Ingelheim, Lilly, Novo Nordisk, Novartis and
468 Sanofi. The remaining authors do not report any competing interests.

- 470 1. Sinha, R. *et al.* Assessment of variation in microbial community amplicon sequencing by the Mi-
471 crobiome Quality Control (MBQC) project consortium. *Nature biotechnology* **35**, 1077–1086;
472 10.1038/nbt.3981 (2017).
- 473 2. Costea, P. I. *et al.* Towards standards for human fecal sample processing in metagenomic studies.
474 *Nature biotechnology* **35**, 1069–1076; 10.1038/nbt.3960 (2017).
- 475 3. Rothschild, D. *et al.* Environment dominates over host genetics in shaping human gut microbiota.
476 *Nature* **555**, 210–215; 10.1038/nature25973 (2018).
- 477 4. Schmidt, T. S. B., Raes, J. & Bork, P. The Human Gut Microbiome: From Association to Modula-
478 tion. *Cell* **172**, 1198–1215; 10.1016/j.cell.2018.02.044 (2018).
- 479 5. Vujkovic-Cvijin, I. *et al.* Host variables confound gut microbiota studies of human disease. *Nature*;
480 10.1038/s41586-020-2881-9 (2020).
- 481 6. Tsuda, A. *et al.* Influence of Proton-Pump Inhibitors on the Luminal Microbiota in the Gastrointes-
482 tinal Tract. *Clinical and translational gastroenterology* **6**, e89; 10.1038/ctg.2015.20 (2015).
- 483 7. Forslund, K. *et al.* Disentangling type 2 diabetes and metformin treatment signatures in the human
484 gut microbiota. *Nature* **528**, 262–266; 10.1038/nature15766 (2015).
- 485 8. Maier, L. *et al.* Extensive impact of non-antibiotic drugs on human gut bacteria. *Nature* **555**, 623–
486 628; 10.1038/nature25979 (2018).
- 487 9. Le Bastard, Q. *et al.* Systematic review: human gut dysbiosis induced by non-antibiotic prescription
488 medications. *Alimentary pharmacology & therapeutics* **47**, 332–345; 10.1111/apt.14451 (2018).
- 489 10. Vich Vila, A. *et al.* Impact of commonly used drugs on the composition and metabolic function of
490 the gut microbiota. *Nature communications* **11**; 10.1038/s41467-019-14177-z (2020).
- 491 11. Vieira-Silva, S. *et al.* Statin therapy is associated with lower prevalence of gut microbiota dysbiosis.
492 *Nature* **581**, 310–315; 10.1038/s41586-020-2269-x (2020).
- 493 12. Falony, G. *et al.* Population-level analysis of gut microbiome variation. *Science (New York, N.Y.)*
494 **352**, 560–564; 10.1126/science.aad3503 (2016).
- 495 13. Jackson, M. A. *et al.* Gut microbiota associations with common diseases and prescription medica-
496 tions in a population-based cohort. *Nature communications* **9**, 2655; 10.1038/s41467-018-05184-7
497 (2018).
- 498 14. Conlon, M. A. & Bird, A. R. The impact of diet and lifestyle on gut microbiota and human health.
499 *Nutrients* **7**, 17–44; 10.3390/nu7010017 (2014).
- 500 15. Blaser, M. J. Antibiotic use and its consequences for the normal microbiome. *Science (New York,*
501 *N.Y.)* **352**, 544–545; 10.1126/science.aad9358 (2016).
- 502 16. Imhann, F. *et al.* Proton pump inhibitors affect the gut microbiome. *Gut* **65**, 740–748; 10.1136/gut-
503 jnl-2015-310376 (2016).
- 504 17. Imhann, F. *et al.* The influence of proton pump inhibitors and other commonly used medication on
505 the gut microbiota. *Gut Microbes* **8**, 351–358; 10.1080/19490976.2017.1284732 (2017).
- 506 18. Wu, H. *et al.* Metformin alters the gut microbiome of individuals with treatment-naïve type 2 dia-
507 betes, contributing to the therapeutic effects of the drug. *Nature medicine* **23**, 850–858;
508 10.1038/nm.4345 (2017).
- 509 19. Koeth, R. A. *et al.* γ -Butyrobetaine is a proatherogenic intermediate in gut microbial metabolism of
510 L-carnitine to TMAO. *Cell metabolism* **20**, 799–812; 10.1016/j.cmet.2014.10.006 (2014).
- 511 20. Kopp, E. & Ghosh, S. Inhibition of NF-kappa B by sodium salicylate and aspirin. *Science (New*
512 *York, N.Y.)* **265**, 956–959; 10.1126/science.8052854 (1994).
- 513 21. Davalli, A. M., Perego, C. & Folli, F. B. The potential role of glutamate in the current diabetes
514 epidemic. *Acta diabetologica* **49**, 167–183; 10.1007/s00592-011-0364-z (2012).

- 515 22. Schmidt, T. S. *et al.* Extensive transmission of microbes along the gastrointestinal tract. *eLife* **8**;
516 10.7554/eLife.42693 (2019).
- 517 23. Shimazu, T. *et al.* Suppression of Oxidative Stress by β -Hydroxybutyrate, an Endogenous Histone
518 Deacetylase Inhibitor. *Science (New York, N.Y.)* **339**, 211–214; 10.1126/science.1227166 (2012).
- 519 24. Shah, B. R. & Hux, J. E. Quantifying the risk of infectious diseases for people with diabetes. *Dia-*
520 *betes care* **26**, 510–513; 10.2337/diacare.26.2.510 (2003).
- 521 25. Korpela, K. & Vos, W. M. de. Antibiotic use in childhood alters the gut microbiota and predisposes
522 to overweight. *Microbial Cell* **3**, 296–298; 10.15698/mic2016.07.514.
- 523 26. Le Chatelier, E. *et al.* Richness of human gut microbiome correlates with metabolic markers. *Nature*
524 **500**, 541–546; 10.1038/nature12506 (2013).
- 525 27. Arumugam, M. *et al.* Enterotypes of the human gut microbiome. *Nature* **473**, 174–180; 10.1038/na-
526 ture09944 (2011).
- 527 28. Costea, P. I. *et al.* Enterotypes in the landscape of gut microbial community composition. *Nature*
528 *microbiology* **3**, 8–16; 10.1038/s41564-017-0072-8 (2018).
- 529 29. Cameron, A. R. *et al.* Anti-Inflammatory Effects of Metformin Irrespective of Diabetes Status. *Cir-*
530 *culation research* **119**, 652–665; 10.1161/CIRCRESAHA.116.308445 (2016).
- 531 30. Dandona, P. *et al.* Increased plasma concentration of macrophage migration inhibitory factor (MIF)
532 and MIF mRNA in mononuclear cells in the obese and the suppressive action of metformin. *The*
533 *Journal of clinical endocrinology and metabolism* **89**, 5043–5047; 10.1210/jc.2004-0436 (2004).



