ARTICLE

OPEN

Check for updates

Imidazole propionate is increased in diabetes and associated with dietary patterns and altered microbial ecology

Antonio Molinaro et al.[#]

Microbiota-host-diet interactions contribute to the development of metabolic diseases. Imidazole propionate is a novel microbially produced metabolite from histidine, which impairs glucose metabolism. Here, we show that subjects with prediabetes and diabetes in the MetaCardis cohort from three European countries have elevated serum imidazole propionate levels. Furthermore, imidazole propionate levels were increased in subjects with low bacterial gene richness and Bacteroides 2 enterotype, which have previously been associated with obesity. The Bacteroides 2 enterotype was also associated with increased abundance of the genes involved in imidazole propionate biosynthesis from dietary histidine. Since patients and controls did not differ in their histidine dietary intake, the elevated levels of imidazole propionate in type 2 diabetes likely reflects altered microbial metabolism of histidine, rather than histidine intake per se. Thus the microbiota may contribute to type 2 diabetes by generating imidazole propionate that can modulate host inflammation and metabolism.

[#]A list of authors and their affiliations appears at the end of the paper.

Type 2 diabetes is a metabolic and societal disease that is associated with an altered gut microbiome[1–6], characterized by a lower abundance of butyrate-producing bacteria[1,2]. Fecal microbiota transfer experiments in humans have demonstrated that gut microbiota can directly affect insulin sensitivity providing causal evidence that gut microbiota can contribute to disease development[7,8]. Diet strongly affects the microbial composition and provides a substrate for microbial enzymes generating metabolites, which can modulate host physiology[9]. Since the microbiome differs between ethnicities and different geographical regions[2,10,11], gut microbiome-derived metabolites might be more conserved biomarkers than specific taxa. Furthermore, metabolites can provide mechanistic insights that may lead to the development of new therapeutic strategies for clinical management of patients with impaired glucose metabolism[12,13].

Accumulating data suggest that microbial metabolism of dietary components contributes to cardiometabolic diseases[14–17], but the full appreciation of the interaction between diet and the microbiome in generating such metabolites is still scarce. Some bacterial metabolites such as secondary bile acids[18], short-chain fatty acids[19], branched-chain amino acids[15], and trimethylamines have attracted significant interest in cardiometabolic diseases[20]. We recently identified that imidazole propionate (ImP) is produced by type 2 diabetes associated microbiome through alternative metabolism of histidine, which induces impaired glucose metabolism by activating the p38γ-mTOR1-S6K1 signaling[16,21].

Here we examine ImP serum levels in a large European multicentric cohort (MetaCardis), from three different European countries, consisting of subjects with different severity of impaired glucose metabolism demonstrating that ImP is increased in subjects with prediabetes and type 2 diabetes. Furthermore, we extend previous studies[16,21] to demonstrate that ImP is associated with, inflammation, altered microbiome, dietary habits, but not histidine intake.

Results

Serum ImP is increased in pre- and type 2 diabetes. The patients with type 2 diabetes in the MetaCardis cohort were slightly older, with a higher proportion of non-Caucasian males compared with healthy individuals and subjects with prediabetes. Patients had an impaired metabolic profile [higher body mass index (BMI) and waist/hip ratio, glucose, insulin, HbA1c, and lipid profiles], while there were no significant differences in kidney function (Table 1).

ImP serum levels were significantly higher in subjects with pre- and type 2 diabetes compared with healthy controls (Fig. 1a), with no impact of subject ethnic background (Supplementary Fig. 1a). Similar observations were made in all enrollments centers (Supplementary Fig. 1b). As compared to subjects in the lowest quartile of ImP levels, those in the highest quartile had a significantly higher risk of having prediabetes [odds ratio (OR) 1.75; 95% confidence interval (CI) 1.18–2.57; $P = 0.006$] and type 2 diabetes [OR 2.76, 95%; CI 1.86–4.12, $P < 0.001$; Fig. 1b, c, Supplementary Table 1, after adjusting for traditional risk factors (Model 1: age, gender, BMI, ethnicity) and for kidney function (Model 2: Model 1 + creatinine clearance)].

We next examined associations between ImP levels and markers of glucose and lipid metabolism as well as for surrogates of insulin resistance. ImP correlated positively with fasting HbA1c, glycemia, insulinemia, HOMA-IR, and triglyceride-glucose index and negatively with HOMA-B. These results indicate a link between circulating ImP and impaired glucose metabolism profiles. Importantly, correlations remained significant after adjustment for known traditional risk factors (Model 1:

age, gender, BMI and ethnicity), for kidney function (Model 2: Model 1 + creatinine clearance), and even the presence of type 2 diabetes (Model 3: Model 2 + diabetes status, Fig. 1d, Supplementary Table 2).

To investigate if ImP also was associated with the dynamic assessment of glucose metabolism, we performed further analyses on a subpopulation ($n = 586$) where oral glucose tolerance tests (OGTT) were performed. Subject stratification based on ImP quartiles revealed that elevated ImP levels were associated with increased glucose, insulin, and C-peptide levels 2 hours after OGGT, translating to reduced Stumvoll sensitivity index (Supplementary Table 3).

All together, serum ImP is increased in pre- and type 2 diabetes and associates with markers of impaired glucose metabolism independently of diabetes status.

ImP is associated with diabetes treatment and co-morbidities. Metformin is the first line of treatment for type 2 diabetes and has a profound effect on microbiota composition and function[22,23]. ImP can alter the glucose-lowering effects of metformin treatment[21]. Thus, we evaluated the effects of metformin and other anti-diabetic drugs on ImP levels (Supplementary Fig. 1c). Patients with metformin-treated type 2 diabetes had higher levels of ImP compared with those without any treatment. This could be due to the altered microbiome following metformin treatment[22,23] or reflecting a more severe disease phenotype that required polypharmacy. Indeed, subjects treated with insulin and additional anti-diabetic drugs also had increased ImP levels. When we performed a sub-analysis including only subjects naïve for anti-diabetic treatments, subjects with type 2 diabetes displayed significantly increased levels of ImP [28.1 nM (16.1–59.2) median and interquartile range; $n = 140$] compared to subjects with prediabetes [27.8 nM (17–49.3) median and interquartile range; $n = 654$] or normal glucose tolerance [19.7 nM (13.2–33.9) median and interquartile range; $n = 359$] ($P = 0.028$, linear regression after adjustment for age, gender, BMI, ethnicity, and creatinine clearance).

When we performed a sub-analysis for the presence of cardiovascular diseases (CVD) we could observe that subjects with CVD had significantly increased levels of ImP [36.7 nM (20.5–69.1) median and interquartile range; $n = 390$] compared to the subjects without CVD [25.2 nM (15.2–47.7) median and interquartile range; $n = 1568$] ($P < 0.001$, linear regression after adjustment for age, gender, BMI, ethnicity, diabetes status, and creatinine clearance) which requires further investigation.

ImP serum levels are associated with an altered microbiome. Since low microbiome gene count is associated with obesity, insulin resistance, and dyslipidemia[24], we next investigated if increased circulating ImP levels are associated with an altered microbiome. Thus, the study population was stratified into high and low gene count (threshold: 607,000 genes) and observed that subjects with low gene count had higher circulating ImP levels compared to those with high gene count, independently of diabetes status (Fig. 2a). We observed a significant negative correlation between ImP residuals (ImP levels adjusted for age, gender, BMI, ethnicity, and creatine clearance) and gene count independently of diabetes status. Interestingly, the effect size was even stronger for type 2 diabetes ($R = -0.31$, $P < 0.001$) compared with prediabetes and healthy subjects (Supplementary Fig. 2a).

The human gut microbiome can be separated into community types, also known as enterotypes[25]. We next analyzed if ImP levels were associated with specific enterotypes and observed that subjects with Bacteroides 2 enterotype had significantly increased serum ImP levels compared with other enterotypes (Fig. 2b). This

Table 1 Clinical and biochemical features of the MetaCardis cohort.

	Healthy controls	Prediabetes	Type 2 diabetes	*P*-value
N	539	654	765	–
Center of enrollment (%)				
France (*n* = 835)	21.4	34.0	44.6	<0.001
Germany (*n* = 587)	34.4	19.3	46.3	
Denmark (*n* = 536)	29.5	47.9	22.6	
Age (years)	53 (39–63)	59 (51–65)	62 (55–67)	<0.001
Male (%)	39	52.6	56.3	<0.001
Non-Caucasian ethnicity (%)	7.2	10.2	20.5	<0.001
Body Mass Index (kg/m*2)	25.9 (22.6–38.2)	30.4 (25.8–39.1)	31.9 (28.3–36.7)	<0.001
Waist/hip ratio	1 (1–1.3)	1.1 (1–1.3)	1.1 (1–1.2)	<0.001
Glucose (nM)	5 (4.7–5.3)	5.7 (5.3–6)	7.5 (6.3–8.9)	<0.001
Insulin (mUI)	6.9 (4.4–11.3)	9.9 (6.6–14.5)	12.5 (8.7–21)	<0.001
HbA1c (%)	5.4 (5.2–5.5)	5.8 (5.6–6)	6.8 (6.3–7.6)	<0.001
Triglycerides (mM/l)	1 (0.7–1.4)	1.2 (0.9–1.6)	1.5 (1.1–2.2)	<0.001
Creatinine clearance (ml/min)	87.5 (77.4–100)	86.4 (75–97.7)	86.4 (74.2–100.2)	0.206
Medications (%)				
Any anti-diabetic treatment	0	0	81.7	<0.001
Metformin	0	0	69.8	<0.001
Any lipid-lowering treatment	21.5	36.2	51.5	<0.001
Statins	21.2	33.9	48.1	<0.001

Data are presented as median and interquartile range or as a percentage. *P*-values for continuous variables were calculated using linear regression. *P*-values for categorical variables were calculated using the Fisher test.

enterotype has been linked to low gene richness and low bacterial cell load as well as impaired metabolism[26,27] and pro-inflammatory conditions such as inflammatory bowel disease[28,29]. We next investigated if specific mOTUs were associated with ImP levels using a random forest approach. We observed that *Clostridium bolteae*, *Clostridium symbiosum,* and *Ruminococcus gnavus* were the most important mOTUs positively associated with ImP after adjustment for age, gender, BMI, ethnicity, creatine clearance, and diabetes status (Fig. 2c and Supplementary Table 4). Increased abundance of these bacteria has previously been reported in subjects with metabolic diseases such as type 2 diabetes and prediabetes[1,5] as well as in subjects with inflammatory bowel disease (IBD)[30]. In contrast, other bacteria with anti-inflammatory capacity[31,32], such as *Faecalibacterium prausnitzii* were negatively associated with ImP serum levels. Taken together these data suggest that ImP serum levels are linked to a pro-inflammatory microbiota composition, in agreement with that ImP initially was identified to be increased in subjects with gut inflammation[33].

ImP is associated with systemic inflammation. Next, we explored whether ImP levels were associated with inflammatory serum markers in the population and observed that serum ImP levels were positively correlated with serum markers of inflammation [total leucocytes count, high sensitive C-reactive protein (hs-CRP), interferon gamma-induced protein 10 (IP-10)] after correction for known traditional risk factors (Model 1: age, gender, BMI, ethnicity), for kidney function (Model 2: Model 1 + creatinine clearance), and for the presence of type 2 diabetes (Model 3: Model 2 + diabetes status) (Fig. 2d and Supplementary Table 5). Moreover, by examining a subpopulation (*n* = 439) from the MetaCardis cohort with peripheral lymphocytes characterization[34], we observed a significant negative correlation between ImP levels and circulating mucosal-associated invariant T cells (MAIT, Fig. 2e and Supplementary Table 6), which have innate effector-like qualities defending against microbial infections. Interestingly, the reduction in peripheral MAIT cells has been linked with metabolic diseases and obesity, and with cardiometabolic disease progression in the MetaCardis population[34].

Taken together, low gene count microbiome and Bacteroides 2 enterotype are associated with increased circulating ImP levels that may contribute to type 2 diabetes by promoting low-grade inflammation.

Microbial metabolism of histidine. To gain further understanding of how the microbiota metabolizes histidine to ImP we analyzed the abundance of the *hutH* gene encoding histidine ammonia lyase and of *urdA*, the gene encoding urocanate reductase. A major challenge for assessing enzyme specificity is that several enzymes with homologous sequences may have different substrate specificity. However, *urdA* can be identified based on amino acids in the FAD-binding domains in the active site[16]. *urdA* is a low abundant gene whereas *hutH* is more prevalent and observed in 201 metagenomic species (Supplementary Table 7). After correction for age, gender, BMI, ethnicity, and creatinine clearance both *hutH* and *urdA* abundances were increased according to diabetes status and ImP quartiles (Supplementary Fig. 2b–e). As expected, both *urdA* and *hutH* abundances were increased in the Bacteroides 2 enterotype (Fig. 2f, g). In agreement with the association between the increased abundance of *hutH* and *urdA* with Bacteroides 2 enterotype, we also observed negative correlations between these genes and gene richness (rho = −0.41, *P* < 0.001 and rho = −0.25, *P* < 0.001 for *hutH* and *urdA*, respectively) (Supplementary Fig. 2f, g).

Unhealthy dietary patterns are associated with serum ImP. Histidine is a precursor of ImP, accordingly, we evaluated the daily dietary histidine intake in our study population. Based on food-frequency questionnaire records, we did not observe any significant differences in histidine intake when the population was stratified according to ImP quartiles (Fig. 3a). Accordingly, we did not observe differences in circulating histidine levels in controls and subjects with type 2 diabetes (*n* = 1895; *P* = 0.78). Next, we evaluated the full spectrum of nutrient intake and identified a significant positive correlation between ImP and saturated fat intake (driven by high cheese intake) and negative correlations with fiber and unsaturated fat intake (driven by reduced intake of vegetables and nuts, Fig. 3b). Moreover, we
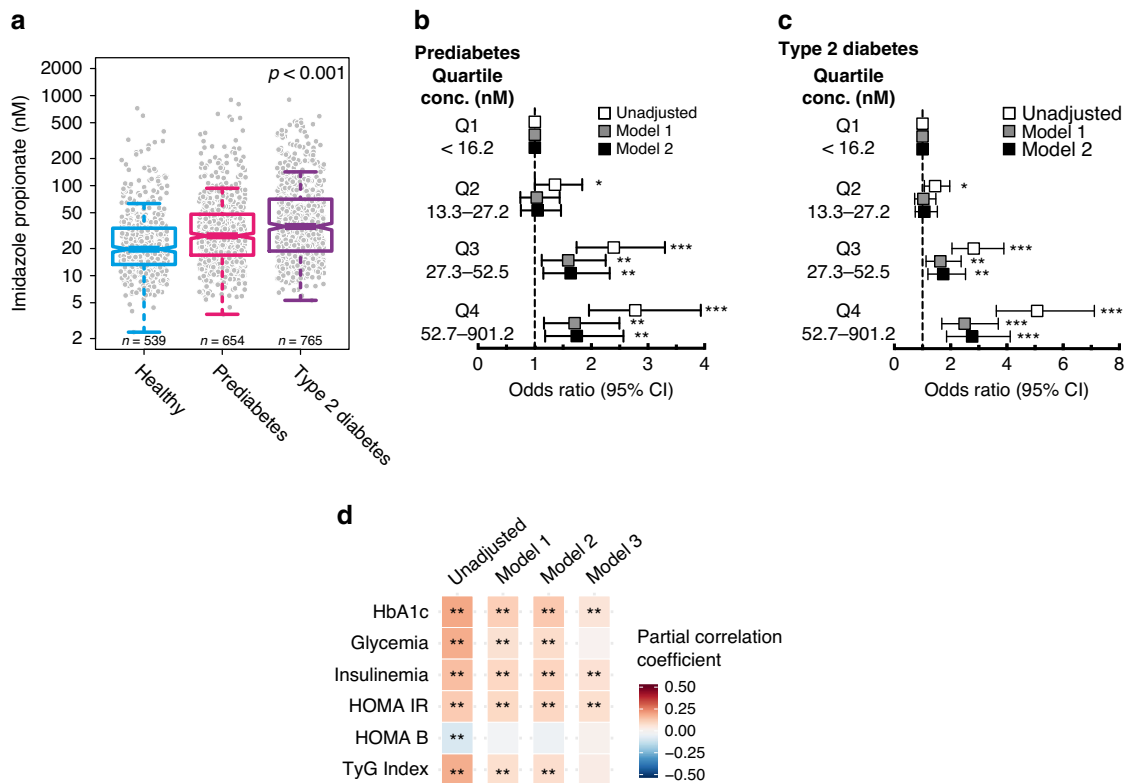
Fig. 1 Imidazole propionate is increased in subjects with type 2 diabetes. a Serum levels of imidazole propionate in healthy subjects ($n = 539$), subjects with prediabetes ($n = 654$), and with type 2 diabetes ($n = 765$). *P*-values were calculated with linear regression. Data are represented as boxplots: middle line is the median, the lower and upper hinges are the first and third quartiles, the upper whisker extends from the hinge to the largest value no further than 1.5× the interquartile range (IQR) from the hinge, and the lower whisker extends from the hinge to the smallest value at most 1.5× IQR of the hinge. Gray dots are single data points. **b**, **c**. Multinomial logistic regression for prediabetes and type 2 diabetes vs. healthy controls according to imidazole propionate quartiles. Odds ratios (OR) were calculated using the lowest quartile of imidazole propionate (Q1) as reference. Model 1 OR was adjusted for age, gender, body mass index (BMI), and ethnicity. Model 2 OR was adjusted for model 1 plus creatinine clearance. Squares represent OR and the upper and lower whisker the 95% confidence intervals (CI), raw data are presented in Supplementary Table 1. *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$ **d**. Correlation matrix for imidazole propionate and glycated hemoglobin (HbA1c), glycemia, insulinemia, homeostatic model assessment of insulin resistance (HOMA-IR), updated HOMA model for beta-cell function (HOMA-B), and the triglyceride and glucose (TyG) index. Pearson partial correlation coefficients and *P*-values were calculated using partial correlations adjusted for Model 1: age, gender, body mass index, and ethnicity. Model 2: Model 1 plus creatinine clearance, Model 3: Model 2 plus diabetes status. False discovery rate (FDR) adjusted *$P < 0.05$, **$P < 0.01$. See also Supplementary Table 2. Source data are provided as a Source Data file.

examined more broadly dietary patterns as assessed by different indexes/scores [the alternate Healthy Eating Index (aHEI), Dietary Approaches to Stop Hypertension (DASH) score, Dietary Diversity Score (DDS) and Mediterranean diet score][35–37]. ImP serum levels correlated negatively with aHEI, DSS, and Mediterranean diet scores, after correction for age, gender, BMI, ethnicity, center (country), daily energy intake, creatinine clearance, and diabetes status. It is important to note that the effect size for these associations is relatively small but overall indicates that an unhealthy diet was associated with increased levels of ImP (Fig. 3b and Supplementary Table 8).

Discussion

The gut microbiome produces a myriad of metabolites that modulate insulin signaling[38], and we recently identified that the histidine metabolite ImP is produced by the microbiome of subjects with type 2 diabetes[16]. Here we demonstrated that ImP is increased in patients with type 2 diabetes in a European multicentric cohort of subjects from different origins. In contrast, to the microbiota that differs between countries[2], regions within a country[11], or even between ethnicities within a city[10], we observed that ImP levels are increased in type 2 diabetes in

French, Danes, and Germans in addition to Dutch[16], Swedes[16], and Chinese[39] as previously reported. Metabolites, such as ImP, may then provide a more relevant indicator of an altered microbial function across populations than metagenomic sequencing per se. Gut microbiome stratification revealed that the Bacteroides 2 enterotype, and reduced gene richness, were associated with increased ImP. Low gene richness has previously been associated with low-grade inflammation, metabolic and inflammatory disorders[24].

Here we observed that ImP was associated with increased levels of pro-inflammatory cytokine and a reduced number of MAIT cells, emanating from the gut mucosa. This is consistent with the fact that ImP is a potent activator of the MAP-kinase p38γ[16]. ImP was originally identified in subjects with gastrointestinal inflammation and is also associated with inflammatory bowel disorders[33]. We have previously demonstrated that type 2 diabetes is associated with increased inflammatory tone in the gut[40], which is consistent with increased ImP levels. Thus one can speculate that the increased levels of ImP in the gut may contribute to elevated inflammatory tone in the mucosal lining.

Here we demonstrate that ImP levels were associated with a low abundance of microbial diversity and Bacteroides 2 enterotype, which previously have been linked to obesity[27] and inflammatory
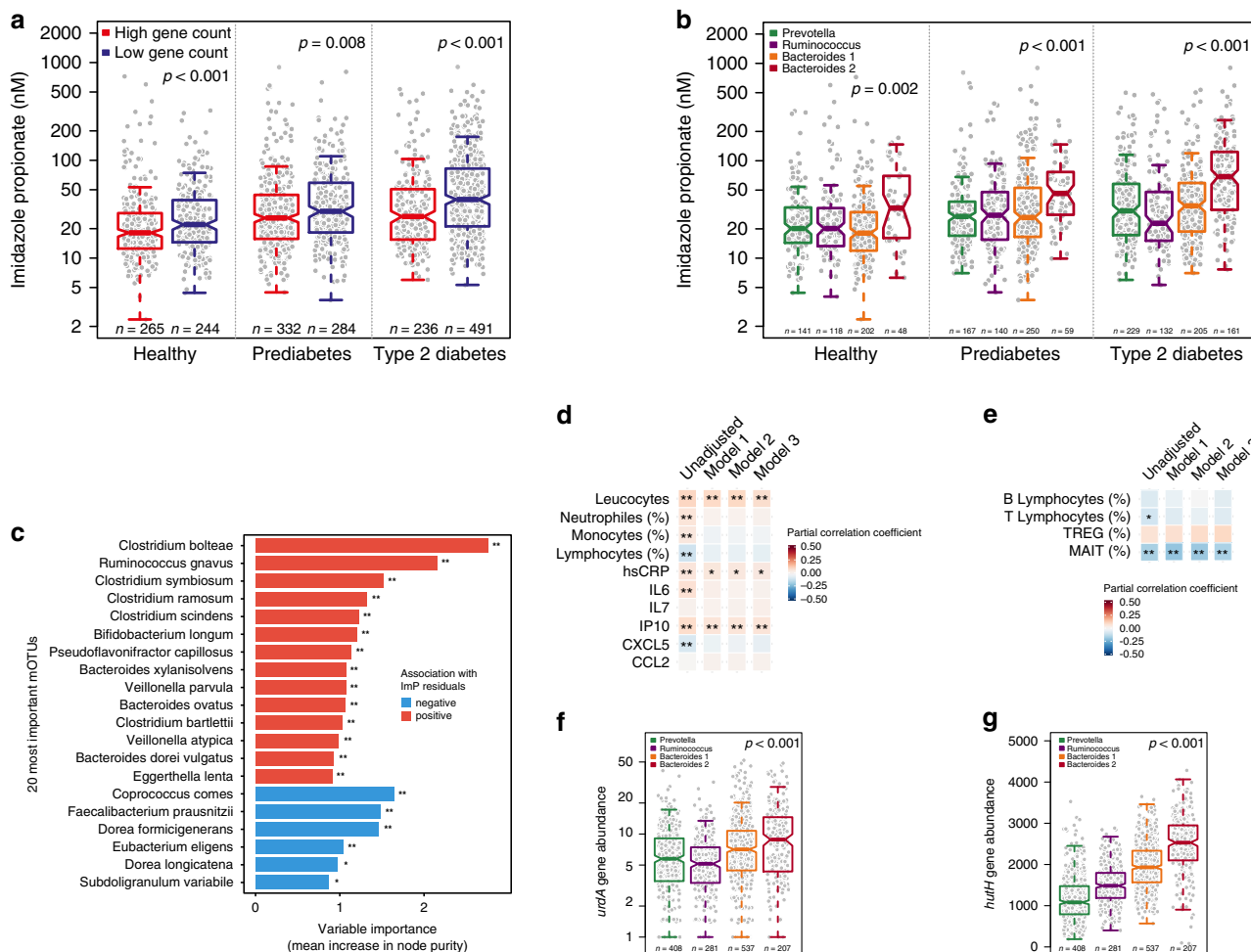
Fig. 2 Imidazole propionate is associated with a pro-inflammatory microbiota. Serum levels of imidazole propionate (ImP) in healthy subjects ($n = 509$), subjects with prediabetes ($n = 616$), and with type 2 diabetes ($n = 727$) according to **a** bacterial gene count and **b** enterotypes. P-values were calculated with linear regression adjusted for age, gender, BMI, ethnicity, and creatinine clearance. **c** Random forest for the 20 most significant mOTUs correlated with ImP residuals, after adjustment for age, gender, BMI, ethnicity, creatinine clearance, and diabetes status. FDR adjusted P-value of spearman correlation between taxa and imp residuals *P < 0.05, **P < 0.01. See also Supplementary Table 4. **d** Partial correlation matrix for ImP serum levels and serum leucocytes count (10^9/l), neutrophils (%), monocytes (%), lymphocytes (%), C-reactive protein (CRP), Interleukin 6 (IL-6), Interleukin 7 (IL-7), Interferon gamma-induced protein 10 (IP-10), C-X-C motif chemokine 5 (CXCL5), chemokine (C-C motif) ligand 2 (CCL2). Pearson partial correlation coefficients and P-values were calculated using partial correlations adjusted for Model 1: age, gender, body mass index, and ethnicity. Model 2: Model 1 plus creatinine clearance, Model 3: Model 2 plus diabetes status. *P < 0.05, **P < 0.01, ***P < 0.001. See also Supplementary Table 5. **e** Partial correlation matrix in a subgroup of patients ($n = 439$) between serum ImP and circulating B- and T lymphocytes (%), regulatory T cells (TREG, %) and mucosal-associated invariant T cell (MAIT, %). Partial correlation coefficients (Pearson for all variables except for MAIT cells for which Spearman coefficient was used since variable distribution remained skewed despite log-transformation) and P-values were calculated using partial correlations for Model 1: age, gender, body mass index, and ethnicity. Model 2: Model 1 plus creatinine clearance, Model 3: Model 2 plus diabetes status. *P < 0.05, **False discovery rate (FDR) adjusted P < 0.05. See also Supplementary Table 6. Relative abundances of *urdA* gene (**f**) and *hutH* (**g**) according to enterotype. P-values were calculated with linear regression adjusted for age, gender, BMI, and ethnicity. For **a**, **b**, **f**, **g** data are represented as boxplots: middle line is the median, the lower and upper hinges are the first and third quartiles, the upper whisker extends from the hinge to the largest value no further than 1.5× the interquartile range (IQR) from the hinge and the lower whisker extends from the hinge to the smallest value at most 1.5× IQR of the hinge. Gray dots are single data points. Source data are provided as a Source Data file.

bowel diseases[28]. These findings are consistent with the fact that subjects with type 2 diabetes have an altered microbiota including reduced diversity and reduced abundance of butyrate-producing bacteria including *F. prausnitzii*[1–6]. Furthermore, recent studies demonstrated that ImP can predict alpha-diversity in humans[13], which also is associated with type 2 diabetes[12]. We observed that an unhealthy diet, but not histidine intake, was associated with increased ImP levels. These findings suggest that rather than being affected by histidine as a substrate, increased ImP production may, at least in part, be the result of an unhealthy diet changing microbial environment and its capacity to produce ImP.

Several reasons could explain the absence of link observed between FFQ-extrapolated histidine intake and ImP levels: (i) ImP production by bacteria requires specific bacterial enzymes[16] and therefore the limiting factor may not be the availability of the substrate (histidine) but the presence of the bacteria with the capacity to produce ImP; (ii) histidine degradation is tightly regulated to maintain sufficient intracellular pools of histidine and Hut enzymes are not formed at maximal rates unless bacteria are limited in other carbon sources most commonly obtained from fiber[41]; (iii) long term dietary habits are key shaping factors for the gut microbiota and an unhealthy diet poor in fiber and
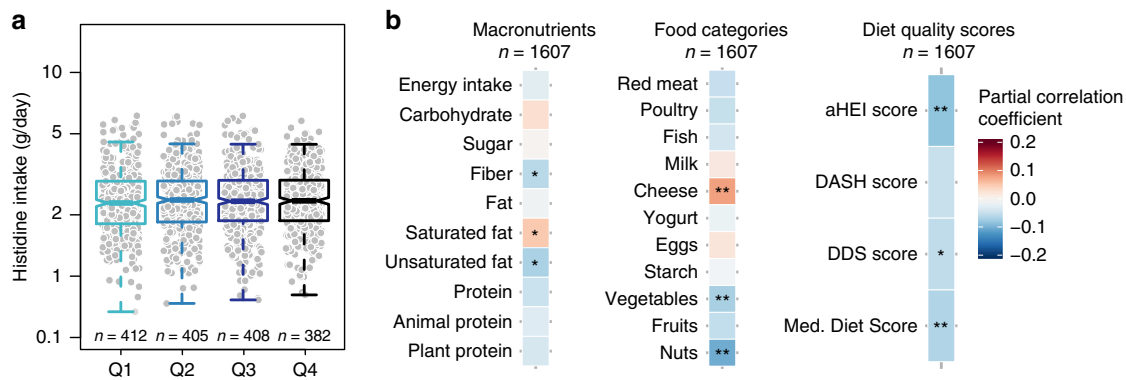
Fig. 3 Imidazole propionate is associated with an unhealthy diet, but not histidine intake. a Dietary histidine intake according to quartiles of imidazole propionate. *P*-values were calculated with linear regression adjusted for age, gender, body mass index, ethnicity, diabetes status, creatinine clearance, daily energy intake (kcal/day), and enrollment center. Data are represented as boxplots: middle line is the median, the lower and upper hinges are the first and third quartiles, the upper whisker extends from the hinge to the largest value no further than 1.5× the interquartile range (IQR) from the hinge, and the lower whisker extends from the hinge to the smallest value at most 1.5× IQR of the hinge. Gray dots are single data points. **b** Correlation matrix for imidazole propionate serum levels and macronutrients, food categories and dietary scores [the alternate Healthy Eating Index (aHEI), Dietary Approaches to Stop Hypertension (DASH) score, dietary diversity score (DDS) and Mediterranean diet score]. Pearson partial correlation coefficients and *P*-values were calculated using partial correlations adjusted for age, gender, body mass index, ethnicity, diabetes status, creatinine clearance, daily energy intake (kcal/day), and enrollment center. *$P < 0.05$, **False discovery rate (FDR) adjusted $P < 0.05$. See Supplementary Table 8. Source data are provided as a Source Data file.

rich in low saturated fats may lead to a dysbiotic microbial environment which will ultimately lead to higher ImP levels[42]; (iv) we cannot exclude that the methods used to capture histidine intake have limited resolution to detect small differences in intake. However, overall, our findings suggest that ImP is not directly linked to dietary histidine intake, but rather an unhealthy diet with reduced intake of fiber and unsaturated fatty acids that results in a dysbiotic microbiome with increased capacity to produce ImP. One limitation of our data is that that we have not used an independent validation cohort to confirm our findings. Further studies using independent cohorts are needed to confirm the role of ImP in type 2 diabetes.

In summary, our data suggest that an unhealthy diet may contribute to an altered microbial community type with increased potential to metabolize dietary histidine to ImP, which in turn contributes to impaired glucose metabolism by activating MAPK signaling leading to degradation of insulin receptor substrate[16] and inflammatory signaling[43]. Since ImP has been observed to be increased in subjects with glucose intolerance and type 2 diabetes of several origins, personalized dietary recommendation or inhibition of *urdA* might be helpful for reducing circulating ImP levels.

Methods

Study population. We examined 1990 subjects from the MetaCardis cohort for whom a serum sample was available but excluded 32 patients due to non-metabolic etiology of cardiovascular diseases ($n = 25$), clear outliers for ImP levels according to Grubb's test ($n = 1$), non-complete biochemistry data ($n = 6$). Subjects were recruited between 2013 and 2015 in clinical institutions in France (Pitié-Salpêtrière Hospital, Center of Research for Clinical Nutrition (CRNH), Institute of Cardio-metabolism And Nutrition (ICAN)), Germany (Integrated Research and Treatment Center (IFB) Adiposity Diseases in Leipzig) and Denmark (Novo Nordisk Foundation Center for Basic Metabolic Research (NNFCBMR) in Copenhagen) for the European project MetaCardis. www.metacardis.net.

Patients with a history of abdominal surgery (other than appendicitis or cholecystectomy), abdominal radiotherapy, digestive cancer or that had received a recent antibiotic treatment (<2 months) were not included. Patients that had participated in the previous cohort-based study were contacted for potential inclusion. A subgroup of healthy control individuals with no signs of obesity or metabolic syndrome were recruited through advertisement and through existing population cohorts. All subjects provided written informed consent and the study was conducted in accordance with the Helsinki Declaration and is registered in clinical trial https://clinicaltrials.gov/show/NCT02059538. The Ethics Committee of each participating country approved the clinical investigation. The study was

approved by the Comite de Protection des Personnes (CPP) Ile de France III no. IDRCB2013-A00189-36.

A detailed list of prescribed medications, anthropometric data, clinical history, fecal sample, and a fasting blood sample was obtained at enrollment. Subjects were classified as healthy, prediabetes, or type 2 diabetes. Type 2 diabetes was defined as fasting glycemia ≥7.0 mmol/l and/or 2 h values during the oral glucose tolerance test >11.1 mmol/l and/or hemoglobin A1c (HbA1c, glycated hemoglobin) ≥6.5% (≥48 mmol/mol) and/or use of any anti-diabetic treatment; prediabetes was defined for subjects without type 2 diabetes as fasting glycemia ≥5.6 mmol/l and/or 2 h values in the oral glucose tolerance test ≥7.8 mmol/l and/or hemoglobin A1c (HbA1c, glycated hemoglobin) ≥5.7% (≥39 mmol/mol) according to the American Diabetes Association (ADA) definitions[44].

Dietary intake data and diet quality assessment. Dietary data were collected via a food-frequency questionnaire that was adapted to the cultural habits of each of the countries of recruitment. A validation study against repeated 24 h-dietary records among 324 French MetaCardis participants has indicated an acceptable validity[45]. Specifically for this study, histidine intake was calculated based on values concerning the histidine content of selected foods published online in the United States Department of Agriculture (USDA) food composition databases https://ndb.nal.usda.gov/ndb/nutrients/. Food groups were further refined by subdividing the original 22 groups into 37, which were used to calculate the total dietary intake of histidine. Dietary quality scores have been adapted from the scores used in the framework of the multicenter European study EPIC[46–49].

For each subject, the basal metabolic rate (BMR) was estimated using Harris and Benedict Formula[50]. Subjects with aberrant energy intake declarations defined as <0.5*BMR or >3.5*BMR were excluded from all nutritional analysis (<10% of the subjects with available nutritional data). In total, 1607 subjects were included in the nutritional analysis.

Biochemical analyses. Blood samples were collected after an overnight fast. Fasting serum glucose, triglycerides, and HbA1c were measured using enzymatic methods. Fasting serum insulin and C-peptide were measured using a chemilu-minescence assay (Insulin Architect, Abbott). High-sensitivity C-reactive protein (hs-CRP) was measured using an IMMAGE automatic immunoassay system (Beckman-Coulter) and high-sensitivity interleukin 6 (hs-IL-6) was measured using the Human IL-6 Quantikine HS ELISA Kit (R&D Systems). IFN-γ–induced protein 10 (IP-10), interleukin 7 (IL-7), C-X-C motif chemokine ligand 2 (CXCL2), and 5 (CXCL5) were measured by using a Luminex assay (ProcartaPlex Mix&-Match Human 13-plex; eBioscience, San Diego, CA, USA).

Oral glucose tolerance test. For 586 subjects of the Metacardis cohort without any clinical/laboratory sign of type 2 diabetes and thus naïve of anti-diabetic treatments, at the inclusion visit, an oral 75 g-glucose tolerance test (OGTT) was performed following standard of care. Serum glucose, insulin, and C-peptide were measured at baseline and 120 min after the glucose load.

Imidazole propionate serum measurements. ImP was quantified using ultra-performance liquid chromatography coupled to tandem mass spectrometry according to previous work. Briefly, serum samples were extracted with 3 volumes of ice-cold acetonitrile containing internal standards (13C3-labeled ImP and urocanate). After derivatization to butyl esters using 5% hydrochloric acid in butanol, the samples were separated on a C18 column using a gradient consisting of water and acetonitrile. Quantification was made using an external calibration curve[16].

Flow cytometry analysis. To characterize and quantify immune cells, 100 μl of whole blood was freshly obtained in a subgroup of subjects ($n = 439$) belonging to one of the centers involved in the MetaCardis consortium (France, Pitié-Salpêtrière Hospital, Institute of Cardiometabolism And Nutrition cytometry platform)[34]. Briefly, blood was incubated with FcR blocking reagent (Miltenyi Biotec, Bergisch Gladbach, Germany), red blood cell lysated, and then white blood cells were stained with the following antibodies: Vioblue-anti-human CD3 (clone BW264/56), PE-Vio770-anti-human CD4 (VIT4), PerCP-Vio700-anti-human CD8 (BW135/80), APC-anti-human CD25 (4E3), FITC-anti-human CD127 (MB15-18C9), PE-anti-human CD161 (191B8) from Miltenyi Biotec, and APC-Cy7-anti-human TCR Vα7.2 (3C10) from BioLegend (San Diego, CA, USA). Data acquisition was performed with a MacsQuant Analyzer using MacsQuantify software (Miltenyi Biotec). Data obtained were analyzed with FlowJo 10.1r5 software (Tree Star, Ashland, OR, USA).

Extraction of fecal genomic DNA and whole-genome shotgun sequencing. Participants collected fecal samples within 24 h before each visit. Samples were either stored immediately at −80 °C or briefly conserved in home freezers, before transport to the laboratory where they were immediately frozen at −80 °C following guidelines[51]. Total fecal DNA was extracted following the International Human Microbiome Standards (IHMS) guidelines (SOP 07 V2 H) and sequenced using ion-proton technology (ThermoFisher Scientific) resulting in 23.3 ± 4.0 million (mean ± SD) 150-bp single-end reads per sample on average. Reads were cleaned using Alien Trimmer (v0.2.4)[39] in order to remove resilient sequencing adapters and to trim low-quality nucleotides at the 3′ side (quality and length cut-off of 20 and 45 bp, respectively). Cleaned reads were subsequently filtered from human and potential food contaminant DNA (using human genome RCh37-p10, *Bos taurus*, and *Arabidopsis thaliana* with an identity score threshold of 97%). The reads were mapped to the Integrated Gene Catalog (IGC) of 9.9 million genes[52], with Bowtie 2.2.4. For each read, the best alignment is conserved. Reads mapped to the main reference with at least 95% of identity are conserved for the counting step if they are not mapped against contaminant references with at least 97% of identity. Gene counts were generated using a two-step procedure (called smart shared counting). First, the unique mapped reads (reads mapping to a unique gene from the catalog) were attributed to the corresponding genes. Second, the shared reads (mapping different genes of the catalog) were attributed according to the ratio of their unique mapping counts. Gene abundance tables (built from mapping against the 9.9 M gene catalog) were processed for richness calculation, downsizing, and normalization using the momr R package. In order to reduce technical bias due to variable sequencing depth, Ion-Proton samples were downsized to 10 million reads, and downsized gene abundances were normalized according to Fragments Per Kilobase per Million mapped reads (FPKM) strategy.

Assessment of gut microbiota characteristics. Metagenomic data were available for 1852 subjects. Abundance for each MGS (metagenomic species) was computed as the mean value of the 50 genes defining a robust centroid of the cluster (if more than 10% of these genes gave positive signals) as proposed[53] for MGS with >500 genes using momr R package. MGS taxonomical annotation was performed using all genes by sequence similarity using NCBI blastN; a species-level assignment was given if >50% of the genes matched the same reference genome of the NCBI database (November 2016 version) at a threshold of 95% of identity and 90% of gene length coverage. The remaining MGS were assigned to a given taxonomical level from the genus to superkingdom if more than 50% of their genes had the same level of assignment. Microbial gene richness (gene count) was calculated by counting the number of genes that were detected at least once in a given sample, using the average number of genes counted in ten independent rarefaction experiments. Alpha-diversity was measured as gene richness i.e., the average number of genes (meaning at least one read mapped) per sample, and subjects were classified in metagenomic richness status (low or high gene count). Metagenomic richness status defined was defined using the threshold separating the bimodal distributions of gene richness in the healthy control group of the German center (where bimodality is revealed) i.e., <607,000 genes for the low gene count group and ≥607,000 genes for the high gene count group. Enterotyping of the cohort was performed following the Dirichlet Multinomial Mixture (DMM) method using microbial taxons (mOTU) abundance matrix of the entire cohort collapsed at the genus level[29,54,55].

hutH analyses. To determine the abundance of *hutH*, the genes encoding histidine ammonia-lyase that metabolizes histidine to urocanate, we quantified the abundance of Kyoto Encyclopedia of Genes Genomes (KEGG) ortholog (KO) K01745.

To identify the metagenomic species with the functional capacity to degrade histidine to urocanate, we projected K01745 on all MGS > 500 genes.

urdA analyses. To determine the DNA abundance of *urdA*, the gene encoding urocanate reductase responsible for ImP production, we used the same pipeline as previously[16] but extend to the latest release of NCBI bacterial genomes which contains 557,951,640 protein-coding genes (as accessed in February 2019) and identified 63,961 potential *urdA* homologs. Exact read mapping was restricted to gene regions containing only 90 nt both down- and up-stream of the FAD active sites of the *urdA* genes based on bowtie 2 with no mismatches and gap opening during reads alignment extracting 12,319 non-redundant sequences for further metagenomic mapping. Of those 12,319 gene regions. In addition, only samples with at least 10 reads mapped and reads with a mapping length larger than 100 bp were included for analysis to ensure that the matched reads cover the active sites. The total reads mapped were then normalized by the sequencing depth in each sample to separate sequences based on true and false urocanate reductases based on the amino acid in position 373. 4760 reads identified as true UrdA without histidine in the 373 position as well as 4968 as with histidine and that position. We have previously shown that histidine in that position prevents urocanate reductase activity[16].

Statistical analysis. The updated homeostatic model assessment of insulin resistance, sensitivity, and beta-cell function (HOMA2-IR, HOMA2-S, HOMA2-B), the quantitative insulin sensitivity check index (QUICKI), the triglycerides and glucose index (TyG) and the Stumvoll index were calculated as described previously[56–60]. Estimated glomerular filtration rate (eGFR) was calculated using the Modification of Diet in Renal Disease (MDRD) formula[61]. All nutrient and food group data are expressed as g of intake per day. Diet quality assessed using four previously validated nutritional scores: the alternate Healthy Eating Index (aHEI), Dietary Approaches to Stop Hypertension (DASH) score, dietary diversity score (DDS), and the Mediterranean diet score[35–37,46]. For descriptive statistics, continuous variables were presented in the median and interquartile range. ImP levels were categorized into quartiles in the full analysis cohort of 1958 individuals. Categorical variables were presented as numbers and percent. Analyses were performed using linear regression models as unadjusted or adjusted for confounding factors (see text and figure legends). For nutritional data, a further adjustment was performed when precised on daily total energy intake and center. In quantitative analysis (graphically presented as heatmaps), partial correlations were performed unadjusted or adjusted for confounders. Variables with skewed distributions were logarithmically transformed before entering the models (age, BMI, creatinine clearance, 2 h glucose, insulin, C-peptide, Stumvoll sensitivity index and Imidazole propionate were log-transformed, Shapiro–Wilk test $P < 0.05$). Correlation analysis of gene richness and functional features (*urdA*, *hutH*) were assessed by Spearman's correlations. The most important mOTUs for the prediction of ImP levels were identified with cross-validated random forest models using ImP residuals adjusted for age, gender, BMI, diabetes status, ethnicity, and creatinine clearance with the randomForest R package. An optimal mtry and max node of 50 for the trees was determined using the mean squared error of ImP residuals in test samples as the outcome. The mean variable importance (using an increase in node purity) was determined with 100-fold cross-validation.

Statistical analyses were carried out using R statistical analysis software version 3.3.2 (http://www.R-project.org/).

Data availability

References

1. Qin, J. et al. A metagenome-wide association study of gut microbiota in type 2 diabetes. *Nature* **490**, 55–60 (2012).
2. Karlsson, F. H. et al. Gut metagenome in European women with normal, impaired and diabetic glucose control. *Nature* **498**, 99–103 (2013).
3. Zhou, W. et al. Longitudinal multi-omics of host-microbe dynamics in prediabetes. *Nature* **569**, 663–671 (2019).
4. Larsen, N. et al. Gut microbiota in human adults with type 2 diabetes differs from non-diabetic adults. *PLoS ONE* **5**, e9085 (2010).

5. Allin, K. H. et al. Aberrant intestinal microbiota in individuals with prediabetes. *Diabetologia* **61**, 810–820 (2018).

6. Wu, H. et al. The gut microbiota in prediabetes and diabetes: a population-based cross-sectional study. *Cell Metab.* https://doi.org/10.1016/j.cmet.2020.06.011 (2020).

7. Vrieze, A. et al. Transfer of intestinal microbiota from lean donors increases insulin sensitivity in individuals with metabolic syndrome. *Gastroenterology* **143**, 913–916 e917 (2012).

8. Kootte, R. S. et al. Improvement of insulin sensitivity after lean donor feces in metabolic syndrome is driven by baseline intestinal microbiota composition. *Cell Metab.* **26**, 611–619.e616 (2017).

9. Schroeder, B. O. & Backhed, F. Signals from the gut microbiota to distant organs in physiology and disease. *Nat. Med.* **22**, 1079–1089 (2016).

10. Deschasaux, M. et al. Depicting the composition of gut microbiota in a population with varied ethnic origins but shared geography. *Nat. Med.* **24**, 1526–1531 (2018).

11. He, Y. et al. Regional variation limits applications of healthy gut microbiome reference ranges and disease models. *Nat. Med.* **24**, 1532–1535 (2018).

12. Menni, C. et al. Serum metabolites reflecting gut microbiome alpha diversity predict type 2 diabetes. *Gut Microbes* **11**, 1632–1642 (2020).

13. Wilmanski, T. et al. Blood metabolome predicts gut microbiome α-diversity in humans. *Nat. Biotechnol.* **37**, 1217–1228 (2019).

14. Pedersen, H. K. et al. Human gut microbes impact host serum metabolome and insulin sensitivity. *Nature* **535**, 376–381 (2016).

15. Newgard, C. B. et al. A branched-chain amino acid-related metabolic signature that differentiates obese and lean humans and contributes to insulin resistance. *Cell Metab.* **9**, 311–326 (2009).

16. Koh, A. et al. Microbially produced imidazole propionate impairs insulin signaling through mTORC1. *Cell* **175**, 947–961.e917 (2018).

17. Tang, W. H. et al. Intestinal microbial metabolism of phosphatidylcholine and cardiovascular risk. *N. Engl. J. Med.* **368**, 1575–1584 (2013).

18. Wahlstrom, A., Sayin, S. I., Marschall, H. U. & Backhed, F. Intestinal crosstalk between bile acids and microbiota and its impact on host metabolism. *Cell Metab.* **24**, 41–50 (2016).

19. Koh, A., De Vadder, F., Kovatcheva-Datchary, P. & Backhed, F. From dietary fiber to host physiology: short-chain fatty acids as key bacterial metabolites. *Cell* **165**, 1332–1345 (2016).

20. Tang, W. H. W., Bäckhed, F., Landmesser, U. & Hazen, S. L. Intestinal microbiota in cardiovascular health and disease: JACC state-of-the-art review. *J. Am. Coll. Cardiol.* **73**, 2089–2105 (2019).

21. Koh, A. et al. Microbial imidazole propionate affects responses to metformin through p38γ-dependent inhibitory AMPK phosphorylation. *Cell Metab.* https://doi.org/10.1016/j.cmet.2020.07.012 (2020).

22. Wu, H. et al. Metformin alters the gut microbiome of individuals with treatment-naive type 2 diabetes, contributing to the therapeutic effects of the drug. *Nat. Med.* **23**, 850–858 (2017).

23. Forslund, K. et al. Disentangling type 2 diabetes and metformin treatment signatures in the human gut microbiota. *Nature* **528**, 262–266 (2015).

24. Le Chatelier, E. et al. Richness of human gut microbiome correlates with metabolic markers. *Nature* **500**, 541–546 (2013).

25. Costea, P. I. et al. Enterotypes in the landscape of gut microbial community composition. *Nat. Microbiol.* **3**, 8–16 (2018).

26. Aron-Wisnewsky, J. et al. Major microbiota dysbiosis in severe obesity: fate after bariatric surgery. *Gut* **68**, 70–82 (2019).

27. Vieira-Silva, S. et al. Statin therapy is associated with lower prevalence of gut microbiota dysbiosis. *Nature* **581**, 310–315 (2020).

28. Vieira-Silva, S. et al. Quantitative microbiome profiling disentangles inflammation- and bile duct obstruction-associated microbiota alterations across PSC/IBD diagnoses. *Nat. Microbiol.* https://doi.org/10.1038/s41564-019-0483-9 (2019).

29. Vandeputte, D. et al. Quantitative microbiome profiling links gut community variation to microbial load. *Nature* **551**, 507–511 (2017).

30. Hall, A. B. et al. A novel Ruminococcus gnavus clade enriched in inflammatory bowel disease patients. *Genome Med.* **9**, 103 (2017).

31. Sokol, H. et al. Faecalibacterium prausnitzii is an anti-inflammatory commensal bacterium identified by gut microbiota analysis of Crohn disease patients. *Proc. Natl Acad. Sci. USA* **105**, 16731–16736 (2008).

32. Miquel, S. et al. Faecalibacterium prausnitzii and human intestinal health. *Curr. Opin. Microbiol.* **16**, 255–261 (2013).

33. Van Der Heiden, C., Wadman, S. K., De Bree, P. K. & Wauters, E. A. K. Increased urinary imidazolepropionic acid, n-acetylhistamine and other imidazole compounds in patients with intestinal disorders. *Clin. Chim. Acta* **39**, 201–214 (1972).

34. Touch, S. et al. Mucosal-associated invariant T (MAIT) cells are depleted and prone to apoptosis in cardiometabolic disorders. *FASEB J.* https://doi.org/10.1096/fj.201800052RR (2018).

35. Fung, T. T. et al. Adherence to a DASH-style diet and risk of coronary heart disease and stroke in women. *Arch. Intern. Med.* **168**, 713–720 (2008).

36. McCullough, M. L. et al. Diet quality and major chronic disease risk in men and women: moving toward improved dietary guidance. *Am. J. Clin. Nutr.* **76**, 1261–1271 (2002).

37. Stefler, D. et al. Mediterranean diet score and total and cardiovascular mortality in Eastern Europe: the HAPIEE study. *Eur. J. Nutr.* **56**, 421–429 (2017).

38. Khan, M. T., Nieuwdorp, M. & Backhed, F. Microbial modulation of insulin sensitivity. *Cell Metab.* **20**, 753–760 (2014).

39. Yu, D. et al. Plasma metabolomic profiles in association with type 2 diabetes risk and prevalence in Chinese adults. *Metabolomics* **12**, 3 (2015).

40. Monteiro-Sepulveda, M. et al. Jejunal T cell inflammation in human obesity correlates with decreased enterocyte insulin signaling. *Cell Metab.* **22**, 113–124 (2015).

41. Bender, R. A. Regulation of the histidine utilization (hut) system in bacteria. *Microbiol. Mol. Biol. Rev.* **76**, 565–584 (2012).

42. Wu, G. D. et al. Linking long-term dietary patterns with gut microbial enterotypes. *Science* **334**, 105–108 (2011).

43. Brown, A. E. et al. p38 MAPK activation upregulates proinflammatory pathways in skeletal muscle cells from insulin-resistant type 2 diabetic patients. *Am. J. Physiol. Endocrinol. Metab.* **308**, E63–E70 (2015).

44. Association, A. D. Classification and diagnosis of diabetes. *Diabetes Care* **42**, S13–S28 (2019).

45. Verger, E. O. et al. Dietary Assessment in the MetaCardis Study: Development and Relative Validity of an Online Food Frequency Questionnaire. *J. Acad. Nutr. Diet.* **117**, 878–888 (2017).

46. Jeurnink, S. M. et al. Variety in vegetable and fruit consumption and the risk of gastric and esophageal cancer in the European Prospective Investigation into Cancer and Nutrition. *Int J. Cancer* **131**, E963–E973 (2012).

47. Consortium, T. I. Adherence to predefined dietary patterns and incident type 2 diabetes in European populations: EPIC-InterAct Study. *Diabetologia* **57**, 321–333 (2014).

48. Sacks, F. M. et al. Rationale and design of the Dietary Approaches to Stop Hypertension trial (DASH). A multicenter controlled-feeding study of dietary patterns to lower blood pressure. *Ann. Epidemiol.* **5**, 108–118 (1995).

49. Sofi, F., Macchi, C., Abbate, R., Gensini, G. F. & Casini, A. Mediterranean diet and health status: an updated meta-analysis and a proposal for a literature-based adherence score. *Public Health Nutr.* **17**, 2769–2782 (2014).

50. Harris, J. A. & Benedict, F. G. A biometric study of basal metabolism in man. *Proc. Natl Acad. Sci. USA* **4**, 370–373 (1918).

51. Thomas, V., Clark, J. & Dore, J. Fecal microbiota analysis: an overview of sample collection methods and sequencing strategies. *Future Microbiol.* **10**, 1485–1504 (2015).

52. Li, J. et al. An integrated catalog of reference genes in the human gut microbiome. *Nat. Biotechnol.* **32**, 834–841 (2014).

53. Nielsen, H. B. et al. Identification and assembly of genomes and genetic elements in complex metagenomic samples without using reference genomes. *Nat. Biotechnol.* **32**, 822–828 (2014).

54. Holmes, I., Harris, K. & Quince, C. Dirichlet multinomial mixtures: generative models for microbial metagenomics. *PLoS ONE* **7**, e30126 (2012).

55. Sunagawa, S. et al. Metagenomic species profiling using universal phylogenetic marker genes. *Nat. Methods* **10**, 1196–1199 (2013).

56. Stumvoll, M. et al. Use of the oral glucose tolerance test to assess insulin release and insulin sensitivity. *Diabetes Care* **23**, 295–301 (2000).

57. Matthews, D. R. et al. Homeostasis model assessment: insulin resistance and beta-cell function from fasting plasma glucose and insulin concentrations in man. *Diabetologia* **28**, 412–419 (1985).

58. Katz, A. et al. Quantitative insulin sensitivity check index: a simple, accurate method for assessing insulin sensitivity in humans. *J. Clin. Endocrinol. Metab.* **85**, 2402–2410 (2000).

59. Levy, J. C., Matthews, D. R. & Hermans, M. P. Correct homeostasis model assessment (HOMA) evaluation uses the computer program. *Diabetes Care* **21**, 2191–2192 (1998).

60. Guerrero-Romero, F. et al. The product of triglycerides and glucose, a simple measure of insulin sensitivity. Comparison with the euglycemic-hyperinsulinemic clamp. *J. Clin. Endocrinol. Metab.* **95**, 3347–3351 (2010).

61. Levey, A. S. et al. A more accurate method to estimate glomerular filtration rate from serum creatinine: a new prediction equation. Modification of Diet in Renal Disease Study Group. *Ann. Intern. Med.* **130**, 461–470 (1999).

Acknowledgements

agreement between the Swedish government and the county councils, the ALF-agreement (ALFGBG- 718101), and the NIHR Imperial Biomedical Research Centre (BRC). The clinical study is sponsored by Assistance Publique Hopitaux de Paris. The computations for metagenomics analyses were performed on resources provided by the Swedish National Infrastructure for Computing (SNIC) through Uppsala Multi-disciplinary Center for Advanced Computational Science (UPPMAX). F.B. is Torsten Söderberg Professor in Medicine and recipient of an ERC Consolidator Grant (European Research Council, Consolidator grant 615362-METABASE).

Author contributions

A.M., P.B., K.C., and F.B. conceived and designed the project. A.M. and P.B. analyzed interactions between ImP and clinical and biochemical data. M.He. and P.B. analyzed ImP levels. H.W. analyzed *urdA* abundance, P.B. and E.B. *hutH* abundance. S.A. and K.A. determined dietary analysis and patterns, C.R. performed multiplex analysis of inflammatory markers, S.A. and F.M. performed immune cell sorting technology and analysis. F.A., J.-E.S., J.-M.O., J.-A.W., T.N., K.C., and M.Stu. recruited patients and R.C., J.-A.W., and T.N. contributed to patient investigation and data management. E.B., S.F., E.L.C., G.F., N.P., E.P., I.L., J.N., and S.-V.S. developed databases, analytical pipelines, and performed metagenomics and functional analysis. B.Q. and H.R. contributed to stool sample processing and sequencing all patient stool samples. C.R., S.A., and F.M. performed all measurements of inflammatory cells and systemic markers, and J.-P.B. performed the analytical measures of metabolic variables. J.-D.Z., S.H., M.-E.D., J.R., J.-M.O., O.P., P.B., M.S., and S.D.E. contributed to results' discussion. A.M., P.B., K.C., and F.B wrote the paper. All authors commented and edited the manuscript.

Funding

Open Access funding provided by Gothenburg University Library.

Competing interests

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

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41467-020-19589-w.

Correspondence and requests for materials should be addressed to K.C. or F.B.

Peer review information *Nature Communications* thanks the anonymous reviewer(s) for their contribution to the peer review of this work.

Reprints and permission information is available at http://www.nature.com/reprints

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

Antonio Molinaro[1,2,35], Pierre Bel Lassen[3,4,35], Marcus Henricsson[1], Hao Wu[1], Solia Adriouch[3], Eugeni Belda[3,5], Rima Chakaroun[6], Trine Nielsen[7], Per-Olof Bergh[1], Christine Rouault[3], Sébastien André[3], Florian Marquet[3], Fabrizio Andreelli[3], Joe-Elie Salem[8], Karen Assmann[3], Jean-Philippe Bastard[9], Sofia Forslund[10], Emmanuelle Le Chatelier[11], Gwen Falony[12,13], Nicolas Pons[11], Edi Prifti[5,14], Benoit Quinquis[11], Hugo Roume[11], Sara Vieira-Silva[12,13], Tue H. Hansen[7], Helle Krogh Pedersen[7], Christian Lewinter[7], Nadja B. Sønderskov[7] & The MetaCardis Consortium*, Lars Køber[7], Henrik Vestergaard[7], Torben Hansen[7], Jean-Daniel Zucker[14], Pilar Galan[15], Marc-Emmanuel Dumas[16,17], Jeroen Raes[11,12], Jean-Michel Oppert[4], Ivica Letunic[18], Jens Nielsen[19], Peer Bork[20,21], S. Dusko Ehrlich[11], Michael Stumvoll[6], Oluf Pedersen[7], Judith Aron-Wisneswky[3,4], Karine Clément[3,4,36]✉ & Fredrik Bäckhed[1,7,22,36]✉

[1]Wallenberg Laboratory, Department of Molecular and Clinical Medicine and Sahlgrenska Center for Cardiovascular and Metabolic Research, University of Gothenburg, 413 45 Gothenburg, Sweden. [2]Department of Medicine, Sahlgrenska University Hospital, Gothenburg, Sweden. [3]INSERM, Nutrition and Obesities; Systemic Approaches (NutriOmics), Sorbonne Université, Paris, France. [4]Assistance Publique Hôpitaux de Paris, Pitie-Salpêtrière Hospital, Nutrition department, CRNH Ile de France, Paris, France. [5]Integromics Unit, Institute of Cardiometabolism and Nutrition, 75013 Paris, France. [6]Medical Department III - Endocrinology, Nephrology, Rheumatology, University of Leipzig Medical Center, Leipzig, Germany. [7]Novo Nordisk Foundation Center for Basic Metabolic Research, Faculty of Health and Medical Sciences, University of Copenhagen, Blegdamsvej 3B, 2200 Copenhagen, Denmark. [8]Assistance Publique Hôpitaux de Paris, Clinical Investigation Center Paris East, 75013 Paris, France. [9]Assistance Publique Hôpitaux de Paris, Biochemistry and Hormonology Department, Tenon Hospital, 75020 Paris, France. [10]Experimental and Clinical Research Center, A Cooperation of Charité-Universitätsmedizin and the Max-Delbrück Center, Berlin, Germany. [11]Micalis Institute, INRA, AgroParisTech, Université Paris-Saclay, Paris, France. [12]Laboratory of Molecular Bacteriology, Department of Microbiology and Immunology, Rega Institute, KU Leuven Leuven, Belgium. [13]Center for Microbiology, VIB Leuven, Belgium. [14]Unité de Modélisation Mathématique et Informatique des Systèmes Complexes, UMMISCO, 93143 Bondy, France. [15]Sorbonne Paris Cité Epidemiology and Statistics Research Centre (CRESS), U1153 Inserm, U1125, Inra, Cnam, University of Paris 13, Nutritional Epidemiology Research Team (EREN), 93017 Bobigny, France. [16]Computational and Systems Medicine, Department of Metabolism, Digestion and Reproduction, Faculty of Medicine, Imperial College London, London SW7 2AZ, UK.

[17]Genomic and Environmental Medicine, National Heart & Lung Institute, Faculty of Medicine, Imperial College London, London SW3 6KY, UK. [18]Biobyte Solutions GmbH, Bothestr. 142, 69117 Heidelberg, Germany. [19]Department of Biology and Biological Engineering, Chalmers University of Technology, SE41128 Gothenburg, Sweden. [20]Structural and Computational Biology, European Molecular Biology Laboratory, Heidelberg, Germany. [21]Molecular Medicine Partnership Unit, University of Heidelberg and European Molecular Biology Laboratory, Heidelberg, Germany. [22]Department of Clinical Physiology, Region Västra Götaland, Sahlgrenska University Hospital, Gothenburg, Sweden. [35]These authors contributed equally: Antonio Molinaro, Pierre Bel Lassen. [36]These authors jointly supervised this work: Karine Clément, Fredrik Bäckhed. *A list of authors and their affiliations appears at the end of the paper. ✉email: karine.clement@inserm.fr; fredrik.backhed@wlab.gu.se

The MetaCardis Consortium

Renato Alves[20], Chloe Amouyal[3,23], Ehm Astrid Andersson Galijatovic[7], Olivier Barthelemy[24], Jean-Paul Batisse[24], Magalie Berland[25], Randa Bittar[26], Hervé Blottière[25], Frederic Bosquet[23], Rachid Boubrit[24], Olivier Bourron[23], Mickael Camus[25], Dominique Cassuto[4], Julien Chilloux[16], Cecile Ciangura[4,23], Luis Pedro Coelho[20,27], Jean-Philippe Collet[24], Maria-Carlota Dao[3], Morad Djebbar[24], Angélique Doré[25], Line Engelbrechtsen[3], Soraya Fellahi[9,28], Leopold Fezeu[15], Sebastien Fromentin[25], Philippe Giral[29], Jens Peter Gøtze[30], Agnes Hartemann[23], Jens Juul Holst[7], Serge Hercberg[15], Gerard Helft[24], Malene Hornbak[7], Jean-Sebastien Hulot[31,32,33], Richard Isnard[24], Sophie Jaqueminet[25], Niklas Rye Jørgensen[30], Hanna Julienne[25], Johanne Justesen[7], Judith Kammer[6], Nikolaj Krarup[7], Mathieu Kerneis[24], Jean Khemis[4], Nadja Buus Kristensen[7], Michael Kuhn[20], Véronique Lejard[25], Florence Levenez[25], Lea Lucas-Martini[4], Robin Massey[25], Nicolas Maziers[25], Jonathan Medina-Stamminger[4], Gilles Montalescot[24], Sandrine Moutel[4], Laetitia Pasero Le Pavin[25], Christine Poitou[3,4], Francoise Pousset[24], Laurence Pouzoulet[29], Sebastien Schmidt[20], Lucas Moitinho-Silva[20], Johanne Silvain[24], Nataliya Sokolovska[3], Sothea Touch[3], Mathilde Svendstrup[7], Timothy Swartz [iD][3,5,34], Thierry Vanduyvenboden[25], Camille Vatier[4] & Stefanie Walther[6]

[23]Assistance Publique Hôpitaux de Paris, Diabetes Department, Pitie-Salpêtrière Hospital, Paris, France. [24]Assistance Publique Hôpitaux de Paris, Cardiology Department, Pitie-Salpêtrière Hospital, Paris, France. [25]INRAE, Metagenopolis, Université Paris-Saclay, Jouy en Josas, France. [26]Assistance Publique Hôpitaux de Paris, Pitie-Salpêtrière Hospital, Biochemistry Department of Metabolic Disorders, Paris, France. [27]Institute of Science and Technology for Brain-Inspired Intelligence, Fudan University, Shanghai, China. [28]Centre de Recherche Saint-Antoine, Sorbonne Université-INSERM UMR-S 938, IHU ICAN, Paris, France. [29]Assistance Publique Hôpitaux de Paris, Endocrinology Department, Pitie-Salpêtrière Hospital, Paris, France. [30]Department of Clinical Biochemistry, University of Copenhagen, Rigshospitalet, Copenhagen, Denmark. [31]Assistance Publique Hôpitaux de Paris, Department of Pharmacology, Pitie-Salpêtrière Hospital, NICO Cardio-oncology Program, CIC-1421, INSERM, Sorbonne Université, Paris, France. [32]PARCC, INSERM, Université de Paris, Paris, France. [33]Assistance Publique Hôpitaux de Paris, Hôpital Européen Georges-Pompidou, CIC1418 and DMU CARTE, Paris, France. [34]Integrative Phenomics, Paris, France.