Physical fitness in community dwelling older adults is linked to dietary intake, gut microbiota and metabolomic signatures

Josué L. Castro-Mejía[1*], Bekzod Khakimov[1], Łukasz Krych[1], Jacob Bülow[2], Rasmus L. Bechshøft[2,3], Grith Højfeldt[2], Kenneth H. Mertz[2], Eva Stahl Garne[1,3], Simon R. Schacht[4], Hajar F. Ahmad[1,5], Witold Kot[6], Lars H. Hansen[6], Federico J. A. Perez-Cueto[1], Mads V. Lind[4], Aske J. Lassen[7], Inge Tetens[4], Tenna Jensen[7], Søren Reitelseder[2,3], Astrid P. Jespersen[7], Lars Holm[2,3,8], Søren B. Engelsen[1], Dennis S. Nielsen[1*]

[1] Department of Food Science, University of Copenhagen, 1958 Frederiksberg C, Denmark

[2] Department of Orthopedic Surgery M, Bispebjerg Hospital, 2400 Copenhagen NV, Denmark

[3] Department of Biomedical Sciences, University of Copenhagen, 2200 Copenhagen N, Denmark

[4] Department of Nutrition, Exercise and Sports, University of Copenhagen, 1958 Frederiksberg C, Denmark

[5] Faculty of Industrial Science and Technology, Industrial Biotechnology Program, Universiti Malaysia Pahang, Pahang, Malaysia

[6] Department of Plant and Environmental Sciences, University of Copenhagen, 1871 Frederiksberg C, Denmark

[7] Copenhagen Center for Health Research in the Humanities, The SAXO Institute, University of Copenhagen, 2300 Copenhagen SV, Denmark

[8] School of Sport, Exercise and Rehabilitation Sciences, University of Birmingham, Birmingham, UK

[*] Correspondence to Josué L. Castro-Mejía (jcame@food.ku.dk) and Dennis S. Nielsen (dn@food.ku.dk)

Email addresses:

J.L.C jcame@food.ku.dk
B.K. bzo@food.ku.dk
Ł.K. krych@food.ku.dk
J.B. jacob.buelow.02@regionh.dk
R.L.B. r.bechshoeft@gmail.com
G.H. grithwh@gmail.com
K.H.M. kenneth.hudlebusch.mertz@regionh.dk
E.G. eva.garne@sund.ku.dk
S.R.C simonschacht@nexs.ku.dk
H.F.A. fauzan@food.ku.dk
W.K. wk@plen.ku.dk
L.H.H. lhha@plen.ku.dk
F.J.A.P. apce@food.ku.dk
M.V.L. madslind@nexs.ku.dk
A.J.L. ajlas@hum.ku.dk
I.T. ite@nexs.ku.dk
T.J. tennaje@hum.ku.dk
S.R. s.reitelseder@gmail.com
A.P.J. apj@hum.ku.dk
L.H. L.Holm@bham.ac.uk
S.B.E. se@food.ku.dk
D.S.N. dn@food.ku.dk

Running title:

Physical fitness in older subjects

Keywords: aging, gut microbiota (GM), host metabolome, physical fitness, proinsulin-C-peptide, energy and dietary fiber intake

Abstract

When humans age, changes in body composition arise along with lifestyle-associated disorders influencing fitness and physical decline. Here we provide a comprehensive view of dietary intake, physical activity, gut microbiota (GM) and host metabolome in relation to physical fitness of 207 community dwelling subjects aged +65 years. Stratification on anthropometric/body-composition/physical-performance measurements (ABPm) variables identified two phenotypes (high/low-fitness) clearly linked to dietary intake, physical activity, GM and host metabolome patterns. Strikingly, despite a higher energy intake high-fitness subjects were characterized by leaner bodies and lower fasting proinsulin-C-peptide/blood glucose levels in a mechanism likely driven by higher dietary-fiber intake, physical activity and increased abundance of Bifidobacteriales and Clostridiales species in GM and associated metabolites (i.e. enterolactone). These factors explained 50.1% of the individual variation in physical fitness. We propose that targeting dietary strategies for modulation of GM and host metabolome interactions may allow establishing therapeutic approaches to delay and possibly revert comorbidities of aging.

1. INTRODUCTION

Throughout the course of aging, physical impairment and changes in body composition may arise along with a number of lifestyle-associated disorders influencing physical decline and ultimately frailty (Xue, 2011; Holm et al., 2014). Aging inevitably occurs in all organisms with genetics, epigenetics and environmental exposures (e.g. diet, physical activity) being modulators of the bodily deterioration caused by biological age (Khan et al., 2017). A number of guidelines toward dietary and daily physical activity recommendations are currently available, however, adherence remains a significant challenge (Gopinath et al., 2016). Further, food perception and dietary habits can be strongly altered during the course of life, particularly those traits associated with the loss of appetite (declined senses of smell and taste), occurrence of immune-senescence and deterioration of the gastro-intestinal system (Giezenaar et al., 2016).

During the last decade, the gut microbiota (GM) has been recognized as a signaling hub that integrates dietary habits with genetic and immune signals throughout life (Thaiss et al., 2016; Peters et al., 2019). Many inflammatory and metabolic disorders, such as obesity, diabetes and inflammatory reactions, are linked with GM dysbiosis (Boulangé et al., 2016). Among Irish older subjects frailty has been linked with changing GM signatures (Claesson et al., 2012) and age-related insulin resistance has been found to be regulated by the metabolic activity (e.g. production of short-chain fatty acids – SCFA) of a number of Clostridiales species (e.g. *Clostridium IV*, *Ruminococcus, Saccharofermentans*) and *Akkermansia muciniphila* (Bodogai et al., 2018; Kong et al., 2016; Biagi et al., 2010). Further, low abundance of these bacteria leads to increased leakage of proinflammatory epitopes from the gut to the blood stream (due to leaky gut syndrome) activating monocytes inflammation and subsequently impair insulin signaling in rodents (Bodogai et al., 2018).

It is well-established, that frail older adults are characterized by changed dietary habits and altered GM and metabolic signatures relative to non-frail peers (Claesson et al., 2012; Lustgarten et al., 2014), but whether similar signatures can be identified among non-

frail older adults of different physical capacity has, to the best of our knowledge, not been investigated previously. A few studies have focused on frail individuals showing that a reduced consumption of dietary fiber compromises the GM associated production of SCFA required for maintenance of colonic epithelial cells and regulation of immune and inflammatory responses (Kong et al., 2016; Biagi et al., 2010; Claesson et al., 2012). Likewise, GM signatures were found to correspond with frailty-indexes in a large cohort of older adults, whose GM composition were inherently driven by dietary patterns (Claesson et al., 2012). Moreover, metabolites related to GM metabolism (e.g. p-cresol sulfate, indoxyl sulfate), peroxisome proliferator-activated receptors-alpha activation, and insulin resistance likely influence physical function in physically impaired older adults (Lustgarten et al., 2014).

Understanding how dietary intake and physical activity in non-frail older adults alter the GM–metabolome axis, and ultimately the physical fitness and the risk of functional decline, is of great clinical interest for the affected subjects as well as for the society. Furthermore, identifying key components of such multifactorial processes may open opportunities to therapeutically address and possibly treat and prevent the comorbidities of aging (Khan et al., 2017). Based on this framework, we characterized dietary intake, daily physical activity, GM and host metabolome in order to be able to explain physical fitness of non-frail older subjects. To this end, we included 207 individuals (65+ years old, self-supportive and apparently healthy) recruited through the Counteracting Age-related Loss of skeletal Muscle mass (CALM) study (http://calm.ku.dk) (Bechshøft et al., 2016). Our findings demonstrate that physical fitness and function corresponded to signatures of fasting proinsulin and average blood glucose, and characterized by clear differences in energy and dietary fiber intake, daily physical activity as well as differential abundance of GM members and a number of fecal and plasma metabolites.

2. RESULTS

2.1 Participants inclusion

Two hundred seven individuals were included in this cross-sectional baseline study (Bechshøft et al., 2016). The recruited subjects are population-level representatives of community dwelling, self-supportive and apparently healthy older adults living in the Danish Capital Region with body mass index (BMI) ranging between 18.5 and 37.3 kg/m^2 (Table 1). Detailed inclusion criteria have been described previously (Bechshøft et al., 2016). From each individual, data were obtained on detailed anthropometric, body-composition and physical-performance measurements (ABPm), average daily physical activity, dietary intake and preferences, GM composition, clinical biomarkers, as well as fecal- and plasma-metabolome adding up to 1,232 analyzed features per subject (Figure S1a).

2.2 Stratification of subjects according to physical fitness and activity monitoring

Study participants were stratified based on non-collinear ABPm variables (Table S1; Variance Inflation Factor, VIF < 2, *r*-coefficient < 0.5) into high- and low-physical fitness phenotypes (level of physical capacity). The selected variables included chair-rise test [30s-test]), BMI, and Dual-energy X-ray Absorptiometry (DXA) scans for body composition (given by legs-soft-tissue fat% (LSF%)), determined as described previously in Bechshøft et al., (2016).

For stratification, hierarchical clustering analysis of principal component analysis (HCP-PCA(Husson et al., 2008)) within sexes was used to determine two fitness phenotypes [high (HF) (n=116) and low (LF) (n=91) (Figure 1a, 1b, Table 2)]. All participants out-performed the suggested ranges for frailty according to the chair-rise test (Jones et al., 1999; Guralnik et al., 1994), while LF phenotypes on average had BMI ranges categorized as overweight (WHO, 2000), as well as a greater deposition of fat mass in their legs (Figure 1b, Table 2). Moreover, 4-day activity monitoring (Dowd et al., 2012) showed significant differences ($p < 0.001$) between the two phenotypes. Longer standing periods (Figure 1c; HF

mean: 4.6 ± 1.3, LF mean: 4.2 ± 1.5) and a greater number of steps per day (Figure 1d; HF mean: 11,129 ± 3,861, LF mean: 8,814 ± 3,595) were recorded among HF phenotypes. The habitual daily activity for LF phenotypes was found to be within recommended ranges (taking approximately 7,000-10,000 steps/day (Tudor-Locke et al., 2011)), but the LF subjects were markedly outperformed by the HF subjects (Figure 1d).

2.3 Dietary food intake in relation to fitness-state

Using 3-day weighted food records (3d-WFR)(Schacht et al., 2019), the daily average energy and macronutrients intake for each person were quantified to obtain an overall view on the dietary intake. On average, the energy intake per person was 24.5 ± 7.4 (range of 11.5 – 55.2) Cal·kg body weight^{-1}·day^{-1}. Protein contributed less of the energy intake (18.9% ± 4.1, range 9-36%) compared to the average energy intake of fat (36.7% ± 7.3, 22-64%) and carbohydrates (44.4% ± 7.7, 17-66%) expressed as percentage of total energy intake.

Total energy consumption per kg body weight (Figure 2a) differed significantly (p < 0.001) between phenotypes, with an average daily intake of 29.3 Cal·kg body weight^{-1}·day^{-1} in HF phenotypes vs. 23.1 Cal·kg body weight^{-1}·day^{-1} in LF phenotypes. The higher energy intake among HF subjects was reflected in a larger fraction of energy (expressed as % energy) from carbohydrates (p = 0.01) as compared to that of dietary protein (Figure 2b and Figure S1b). The same pattern was also observed across daily average intake (g·kg body weight^{-1}·day^{-1}) of dietary fiber (p < 0.0001), starch (p < 0.0001), simple sugars (p = 0.0002) and saturated fatty acids (p = 0.0001) (Figure 2c). Moreover, significant (p < 0.0001) negative correlations between BMI with dietary fiber consumption (r = -0.52) (Figure 2d) energy intake (r = -0.52), starch (r = -0.35) and simple sugars (r = -0.35), as well as positive associations between chair-stand test and energy intake (r = 0.25) were found (Figure S1c-f). Questionnaires on food-choices showed that HF subjects to a higher degree than LF subjects consider healthy food as an important element of their daily life (Figure S1g).

A considerable proportion of subjects from both phenotypes did not comply with the recommended minimum proportion of energy obtained from carbohydrates (Figure 2e) and dietary fiber intake (Figure 2f) as established by the Nordic Nutrition Recommendations (Nordic Council of Ministers, 2012). Yet, the frequency of compliers-to-non-compliers was significantly higher (carbohydrates: $p = 0.006$, dietary fiber: $p = 0.03$) in HF individuals. Furthermore, using the Goldberg cut-off (Black, 2000), 46 under-reporters (UR) and 2 over-reporters (OR) of energy intake were identified. Nonetheless, if excluded, individuals with higher physical capability (HF phenotype) still had a higher energy ($p < 0.001$) and energy from carbohydrates ($p < 0.06$) intake as compared to LF subjects (Table S2). Since UR and OR subjects did not change the overall findings they were not excluded in downstream analyses.

2.4 Characterization of GM and correspondence with fitness and diet

Sequencing of DNA extracted from stool samples yielded 11.3 million reads derived from the 16S rRNA-gene V3-region with an average of 116,476 (48,872 SD) sequences per subject. The analysis of amplicon-sequencing data generated 10,084 zOTUs (sequence variants) summarized over 875 cumulative species (species richness) and 8 core-species (defined as being present in all recruited subjects) among the study subjects (Figure S2a). The relative abundance of core species varied between 18 – 84% (Supplementary Figure 2b). Between sexes no significant differences in beta-diversity (Figure 3a) and alpha-diversity (Figure S2c) were observed. Furthermore, regardless of sex, the study participants were characterized by higher relative abundance of e.g. Lachnospiraceae spp., *Akkermansia* spp., *Blautia* spp., along with reduced proportions of *Bacteroides* spp. (Figure S2d) as compared to the community-dwelling group of older adults recruited for the Irish ELDERMET study (Claesson et al., 2012). This may reflect differences associated with dietary habits, age [mean age: baseline-CALM 70 ± 4y, ELDERMET 78 ± 8y], and geographical location.

A substantial higher alpha-diversity ($p = 0.06$, Observed Species) were observed (Figure S2c) among HF phenotypes compared to LF phenotypes, as well as weak but significant ($p < 0.05$) correlations of observed species with BMI, energy and starch intake (Figure S2e-g). Correspondence analysis and analysis of similarities (ANOSIM) on Bray-Curtis (weighted beta-diversity) distance metric calculated from species-level abundance showed significant correspondence ($p = 0.04$) and dissimilarities ($p = 0.01$) in GM composition in connection with the two physical phenotypes (Figure 3b-c).

Also, GM composition was clearly associated with ($p < 0.05$) gradients of energy consumption (Figure 3d), starch (Figure 3e), dietary fiber (Figure 3f) steps per day (Figure 3g) and BMI (Figure 3h) reflecting fitness phenotypes. Using regularized canonical correlation (rCC) analysis associations between those lifestyle covariates (e.g. dietary factors and physical activity) with 161 microbial species were disclosed (Figure 3i, Figure S3) explaining <5% and 13% of the total variance of the microbiota and lifestyle covariates, respectively (Figure S3a). The strongest associations (those > $|0.2|r$, number of species in brackets) were observed for Bacteroidales (12), Bifidobacteriales (2), Clostridiales (106), Coriobacteria (7), Enterobacterales (3), Erysipelotrichales (12), Lactobacillales (3), PAC001057 (Mollicutes members) (8), Proteobacteria (1) and other orders (7) (Figure S3b). Increased intake of energy, starch, dietary fiber, as well as steps per day correlated positively with the relative abundance of up to 103 of those species (e.g. higher Bifidobacteriales abundance) and correlated negatively with BMI (e.g. Proteobacteria being signatures for high BMI) (Figure 3i, Figure S3b).

2.5 Host metabolic state in relation to fitness and dietary intake

Untargeted Gas Chromatography-Mass Spectrometry (GC-MS) metabolomics of human fecal extracts and blood plasma as well as targeted SCFA analysis using GC-MS generated a total of 304 analytes (181 analytes in the fecal and 123 analytes in the plasma metabolome). Nearly half of the metabolites variables were identified, either at level 1 or

level 2 according to the Metabolomics Standards Initiatives (Sumner et al., 2007). These metabolites were monosaccharides, amino acids, organic acids, sterols and long-, and short chain fatty acids. In addition, 31 biomarkers for immunological function, renal and liver function, as well as glucose and lipid metabolism were acquired through blood clinical profiling.

Correspondence analysis on the combined metabolome blocks showed weak discrimination of sexes (Figure 4a) and pronounced discrimination between fitness phenotype (Figure 4b) based on their metabolic profile. Variations in metabolome composition corresponded clearly ($p < 0.05$) with energy intake and consumption of dietary fiber, starch, simple sugars (Figure 4c-f), as well as steps per day and hours-standing-per-day (Figure 4g-h, including stratifying variables: BMI (Figure 4i), chair-stand and LFT%, Figure S4a-b). Likewise, rCC analysis showed significant associations between lifestyle covariates and 34 clinical/metabolic variables (Figure 4j), explaining 9% and 15% of the total variance of the metabolome and lifestyle covariates, respectively (Figure S4c). The strongest associations (> $|0.2|r$) were observed for 19 clinical biomarkers, 10 gut metabolites and 5 plasma metabolites (Figure 4j). Increased intake of energy, starch, dietary fiber (or dietary covariates), as well as steps per day correlated positively with mono- and di-saccharides and negatively with amino acids (Pro, Ala, Trp), glucose metabolism parameters (proinsulin, glucose HbA1c, HbA1c), lipid metabolism (triglycerides, vLDL) and renal function (creatinine, inversely to estimate glomerular filtration rate (eGFR)) measurements, primary bile acids (lithocholic acid) and N-Nitrosotrimethylurea (Figure 4j). Moreover, a higher proportion of enterolactone in the fecal metabolome of HF subjects were also found (Figure 4k). Remarkably, the concentrations of SCFA as well as other/branched-chain fatty acids (O/B-CFA) in the fecal samples did not differ according to phenotypes ($p > 0.13$) or dietary intake factors (Figure 4l-m).

2.6 Dietary intake, gut microbiota and metabolic signatures explain fitness levels independently from physical activity

Characterization of subjects after variable selection based on Random Forest and backward elimination procedure selected 55 variables (Figure 5a) with different levels of importance (Figure 5b) that discriminate the two phenotypes with a high level of accuracy (Figure 5c-d). The features included 25 bacterial species belonging to 7 bacterial orders (Clostridiales, Saccharibacteria, Bacteroidales, PAC001057, Enterobacterales, Erysipelotrichales and Bifidobacteriales), seven dietary components (energy, saturated fatty acids, simple sugars, starch and dietary fiber intake, and energy derived from proteins and carbohydrates); five clinical biomarkers (alanine transaminase, triglycerides, vLDL, fasting proinsulin, average blood glucose/HbA1c). In addition, seven plasma metabolites (amino acids and organic acids), ten fecal metabolites (sugar alcohols, amino acids, primary bile acids and urea) and physical activity (steps per day) were also tabbed (Figure 5a).

Discrimination of the two phenotypes based on all the selected features (combined datasets) had the highest level of accuracy (22% out-of-bag error rate, OOB), followed GM and clinical/metabolome features (23% OOB), dietary intake (36% OOB) and physical activity parameters (46% OOB) (Figure 5d). Through redundancy analysis (RDA) the effect of the selected variables (within blocks) on the stratifying variables it was found that GM had the largest explanatory power (24.7%), followed by dietary intake (17.3%), clinical biomarkers (16.8%), gut metabolome (8.8%), plasma metabolome (6.2%) and physical activity (5.2%) (Figure 5e). Notably, the cumulative explained variance conferred by the pool of selected features reached 50.1%, and even after conditioning the effect of physical activity over the stratifying variables, the cumulative explained variance reached up to 44.9% (Figure 5f).

3. DISCUSSION

The number of older-adults over the age of 65 will increase by more than 50% worldwide over the next three decades (NIH 2011), potentially with huge implications for the health and economy of the implicated individuals and society as a whole. With this, understanding the physical mechanisms and lifestyle conditions linked to fitness and independence in older adults becomes a relevant field of research.

Despite the homogeneity of the recruited subjects (all non-frail and without serious disease) noticeable differences in fitness level was observed and based on non-collinear ABPm variables (chair-rise test, BMI and DXA-scan based body composition), two fitness phenotypes (LF and HF) were identified. Neither of the fitness types were frail (Guralnik et al., 1994) , nevertheless dietary, GM and host metabolome factors were found to clearly discriminate between the two fitness types. HF subjects were characterized by a higher consumption of foods of plant origin as also reflected by their higher levels of total carbohydrates (i.e. starch, simple sugars) and dietary fiber, accompanied by a higher adherence to the recommended intake of carbohydrates and dietary fiber intake given by the Nordic Nutrition Recommendations (Nordic Council of Ministers 2012). These differences were observed in spite of the methodological limitations of 3d-WFR to capturing long term variability (Yang et al., 2010). Furthermore, whether awareness of dietary guidelines influenced the selection of dietary choices in the study participants remains to be investigated, but it is worth mentioning that HF subjects consider healthy food as an important component in their life as also described by Schacht et al., (Schacht et al., 2019).

The GM community and host metabolome clearly discriminated between the HF and LF phenotypes and was largely associated with the consumption of total energy, and plant derived nutrients (such as starch and dietary fibers as well as enterolactone, all being higher in HF subjects). A number of features (Figure 5a) selected from GM, host metabolome, dietary intake and daily physical activity were able to strongly discriminate and explain variation between phenotypes, thereby indicating their strong association with physical function. Daily

physical activity showed the lowest power towards phenotypic differentiation (in spite of the high validity of the method for activity monitoring (Dowd et al., 2012)) and explaining only 5% of the phenotypic variance. Albeit conditioning for physical activity, the remaining set of selected features explained up to 45% of the total variance of the stratifying variables. In particular dietary intake (17% of explained variance), GM composition (24%) and host metabolome (25%) signatures are important drivers of phenotypic differentiation (Figure 5), and also described in animal models (Fujisaka et al., 2018). Accordingly, HF subjects showed a higher proportion of GM members commonly known for their protective roles, such as *Bifidobacterium adolescentis* and *Christensenella* species (Goodrich et al., 2014), and whose abundance corresponded negatively with glucose and lipid metabolism biomarkers (proinsulin, HbA1c, vLDL, triglycerides). Contrarily, LF phenotypes had increased levels of the same biomarkers and a higher relative abundance of pro-inflammatory microbial members in the gut, as for example Enterobacterales (Fei & Zhao 2013; Hoarau et al., 2016; Khan et al., 2014).

SCFAs derived from GM activity have been identified as signaling molecules responsible for maintenance of the integrity of colonic epithelium, glucose homeostasis, lipid metabolism and appetite regulation (Morrison et al., 2016). Claesson et al., (Claesson et al., 2012) reported higher SCFA concentrations (acetate, butyrate and propionate) in the fecal metabolome of older adults living as community-dwellers compared to frail individuals living in residential care. Moreover, decreasing concentrations of these SCFAs were associated with advanced levels of frailty given by diet and specific transitions in GM composition (Claesson et al., 2012). However, in the present study no correlations between fecal SCFA and O/B-CFA concentrations with neither macronutrient distribution or fitness phenotype were found. This suggest that levels of physical function amidst healthy older adults may not be primarily dependent upon changes in the production of these compounds. Instead, this could be due to signals of glucose metabolism deterioration as reflected by significantly ($p < 0.001$) higher proinsulin levels and higher average blood glucose (determined by HbA1c-levels) in the LF

phenotypes (1/116 HF and 20/91 LF subjects had higher than normal ranges of proinsulin (Chi-Squared $p < 0.001$), 10/116 HF and 30/91 LF had higher ranges than those recommended for HbA1c (Gardner & Shoback 2011) (Chi-Squared $p < 0.001$), see Table S3). High concentrations of proinsulin indicates high insulin secretion and hence diminished peripheral insulin sensitivity resulting in a number of metabolic conditions, compromising muscle strength and physical performance (Segerström et al., 2011). Proinsulin was the most important feature of phenotype discrimination and corresponded inversely with the abundance of *Bifidobacterium adolescentis* and several species of *Christensenella,* and Ruminococcaceae (Figure 5a), strongly indicating that GM-proinsulin interactions could be mediators of fitness phenotype. *Bifidobacterium* species (including *B. adolescentis*) have previously been described as promoters of adiponectin and decreasing expression of interleukin-6, both playing prominent roles in metabolic derangements associated with glucose regulation and fatty acid oxidation (Su et al., 2015; Straub & Scherer 2019; Aoki et al., 2017). *Christensenella minuta* (another Clostridiales member) is enriched in individuals with low BMI and has been demonstrated to reduce weight gain and adiposity in mice (Goodrich et al., 2014). Furthermore, while playing a protective role against inflammation, some Clostridiales members act as promoters of regulatory T-cells by interacting with toll-like receptors 2 (TLR2) on intestinal epithelial cells (Kashiwagi et al., 2015). Contrarily, species of Enterobacterales have been consistently linked with insulin resistance and inflammatory responses (Fei & Zhao 2013; Hoarau et al., 2016; Khan et al., 2014), and by means of cell epitopes (i.e. LPS) they interact with TLRs triggering pathogen recognition, low-grade inflammation (Franceschi & Campisi, 2014) and fat accumulation in adipose tissue that ultimately influence muscle strength (Boulangé et al., 2016).

In summary, our findings suggest that dietary patterns underlie mechanisms of physical phenotype differentiation among well-functioning community dwelling older adults, particularly as a driver of GM and glucose metabolism interactions. Despite the limitations of this study related to its inherent cross-sectional nature, the results provide strong evidence

372 emphasizing the central role of diet towards the onset of physical deterioration and its
373 implications prior to clinical manifestations of frailty, e.g. muscle composition and
374 diminished strength (Xue, 2011). Many of the dietary, GM and metabolomic signatures seen
375 in frail older adults (Claesson et al., 2012; Bodogai et al., 2018; Kong et al., 2016; Lustgarten
376 et al., 2014) are already evident in the non-frail, community-dwelling older-adults of low-
377 fitness of this study, pointing at the importance of early intervention strategies, also in this
378 age group. Thus, in view of these findings, developing strategies to improve awareness and
379 adherence to dietary recommendations (complying with dietary reference intakes or even with
380 personalized nutrition (Zeevi et al., 2015)), targeting the regulation of GM and host
381 metabolome interactions, can open opportunities to delay the comorbidities of aging.

382 ## 4. EXPERIMENTAL PROCEDURES

383 ### 4.1 Study Participants

384 Procedures of the CALM project (Clinical Trials NCT02115698) were approved by
385 the Danish Regional Committees of the Capital Region (H-4-2013-070), performed according
386 to the Declaration of Helsinki II and the experimental designed followed as previously
387 described (Bechshøft et al., 2016). For the current study, two hundred and seven subjects (65+
388 years of age) were selected at baseline of the CALM intervention project following the
389 criteria described in Bechshøft et al.,(Bechshøft et al., 2016). Participants were not allowed to
390 take part in any organized sports or resistance training more than once a week, did not suffer
391 from defined metabolic-, tissue-, or gastro-intestinal disorders, nor were prescribed antibiotics
392 3 months prior sample collection and enrollment.

393 ### 4.2 Samples and metadata collection

394 At baseline, participants completed a 3-day weighted food record where food and
395 beverage intake were registered for 3-consecutive days (Wednesday to Friday). Dietary
396 information was typed into the electronic dietary assessment tool, VITAKOST™ (MADLOG

APS, Kolding, Denmark), which uses the Danish Food Composition Databank (version 7.01; Søborg; Denmark) to estimate individual energy and macronutrient intake.

Fecal and blood plasma samples were collected and handled according to the following procedures: (i) fecal samples were kept at 4°C for maximum 48 h after voidance, and stored at -60°C until further use; (ii) overnight-fasted-state (OFS) plasma-samples were collected and deposited in heparin, centrifuged at 3,000×g for 10 min at 4°C, and then stored at -60°C.

For screening of blood-biomarkers, the following tests were performed: complete blood count (CBC), proinsulin-C-peptide (P-CP), glycosylated hemoglobin (HbA1c), coagulation factor, estimate glomerular filtration rate (eGFR), thyroid-stimulating hormone (TSH), and iron-ferritin test determined as described in Bechshøft et al., (Bechshøft et al., 2016) For anthropometric and functional capacities, height (cm) and body-weight (kg) in OFS were measured. Average fast-pace gait speed was measured on an indoor 400 m horizontal track. Number of chair-stands in 30s from a standard table chair was recorded. Relative legs-soft-tissue fat% (LSF%) was determined as an estimate of legs-soft-tissue fat-free and fat-mass based on a dual energy x-ray absorptiometry (DXA) scan (Lunar iDXA Forma with enCORE Software Platform version 15, GE Medical Systems Ultrasound & Primary Care Diagnostics, Madison, WI, USA) performed on participants in overnight fasted state.

4.3 Quantitative questionnaires on food habits

Quantitative questionnaires contained information on food habits, perceptions and preferences, as well as information about life style changes and dietary habits over the life course (Bechshøft et al., 2016).

4.4 GM and metabolomics

Procedures for profiling and process GM and metabolomics data are described in Supplementary Methods.

4.5 Statistical Analyses

Stratification of individuals was based on ABP measurements using the variables described in Table S1. Collinear variables were initially removed, leaving chair-stand [30s-test]), DXA scans (legs-soft-tissue fat% determined in both legs) and BMI as features with a variance inflation factor (VIF) < 2 and *r*-coefficient < 0.5. Subjects were divided according to sex, and a hierarchical clustering analysis of principal component analysis (Husson et al., 2008) was performed on the selected variables (100 iterations).

For univariate data analyses, pairwise comparisons were carried out with unpaired two-tailed Student's *t*-test, Pearson's coefficient was used for determining correlations and Chi-Square test for evaluating groups distributions. For multivariate data analyses, the influence of covariates (e.g. dietary components, BMI, etc.) on data blocks (GM and metabolome) were assessed with (Constrained-) Correspondence Analysis with permutation tests (1,000 permutations), as well as analysis of similarities (ANOSIM test, 999 permutations) on Bray-Curtis distances (implemented in the *Vegan* R-package (Oksanen et al., 2015)).

Correlation of covariates with the same datasets were determined with regularized canonical correlation (rCC) analysis using the *mixOmics* R-package (González et al., 2012). rCC was crossed-validated (leave-one-out approach) with grids (lambda 1 & 2) of 0.05 to 1.0 and a length of 20.

Feature selection for combined datasets was performed with Random Forest. Dataset was randomly divided 200x (200 subsets) into training (70%) and test sets (30%), keeping this proportion over the number of subjects within each fitness group for every split. For a given training set, the *party* R-package (Hothorn et al., 2016) was run for feature selection using unbiased-trees (cforest_unbiased with 6,000 trees) and AUC-based variable (varimpAUC with 100 permutations), and subsequently the selected variables were used to predict (6,000 trees with 1,000 permutations) their corresponding test set using *randomForest* R-package (Liaw & Wiener 2014). The features derived from the subset with a prediction rate

within 1 SD above the mean-prediction (based on the 200 subsets) were selected and subsequently, subjected to sequential rounds of feature selection (following the same tuning of unbiased-trees and AUC-based variable) until prediction could no longer improved. Variation partitioning of stratifying variables (BMI, CS and LSF%) based on selected features derived from the different datasets (i.e. GM, diet, host-metabolome, physical activity) was performed using redundancy analysis (RDA) (Oksanen et al., 2015).

ACKNOWLEDGEMENTS

This project was supported by the University of Copenhagen-funded project "Counteracting Age-related Loss of Skeletal Muscle (CALM)", the Innovation Foundation Denmark-funded project COUNTERSTRIKE (4105-00015B), stipends from the University of Copenhagen, as well as a PhD-stipend from Universiti Malaysia Pahang, Malaysia and Ministry of Education, Malaysia.

AUTHOR CONTRIBUTIONS

Conceptualization: D.S.N., J.L.C., S.B.E., L.H., A.P.J.; Methodology: D.S.N., J.L.C., S.B.E., L.H., A.P.J., A.J.L., T.J., S.R., R.L.B.; Formal Analysis: J.L.C., B.K., Ł.K., D.S.N., S.B.E., L.H; Writing – Original Draft: J.L.C., D.S.N.; Investigation, review & editing: all authors; Visualization: J.L.C., D.S.N.; Supervision: D.S.N., S.B.E., L.H.; Funding Acquisition: D.S.N., S.B.E., L.H., A.P.J.

CONFLICT OF INTEREST

None declared.

Data Availability

Sequence data is available at the European Nucleotide Archive, accession number ENA: PRJEB33008 ([dataset] Castro-Mejía et al., 2019). The remaining data that support the findings of this study are available on request from the corresponding authors. The data are not publicly available due to privacy or ethical restrictions.

REFERENCES

Aoki, R., Kamikado, K., Suda, W., Takii, H., Mikami, Y., Suganuma, N., … Koga, Y. (2017). A proliferative probiotic Bifidobacterium strain in the gut ameliorates progression of metabolic disorders via microbiota modulation and acetate elevation. *Scientific Reports,* 7, 43522. http://dx.doi.org/10.1038/srep43522

Bechshøft, R.L., Reitelseder, S., Højfeldt, G., Castro-Mejía, J.L., Khakimov, B., Ahmad, H.F., … Holm L. (2016). Counteracting Age-related Loss of Skeletal Muscle Mass: a clinical and ethnological trial on the role of protein supplementation and training load (CALM Intervention Study): study protocol for a randomized controlled trial. *Trials* 17, 397. https://doi.org/10.1186/s13063-016-1512-0

Biagi, E., Nylund, L., Candela, M., Ostan, R., Bucci, L., Pini, E., … De Vos, W. (2010). Through ageing, and beyond: Gut microbiota and inflammatory status in seniors and centenarians. *PLoS One,* 5, e10667, https://doi.org/10.1371/journal.pone.0010667

Black, A.E. (2000). Critical evaluation of energy intake using the Goldberg cut-off for energy intake : basal metabolic rate . A practical guide to its calculation ., use and limitations. *International Journal of Obesity,* 24, 1119–1130.

Bodogai, M., O'Connell, J., Kim, K., Kim, Y., Moritoh, K., Chen, C., … Biragyn, A. (2018). Commensal bacteria contribute to insulin resistance in aging by activating innate B1a cells. *Science Translational Medicine,* 10, eaat4271. https://doi.org/10.1126/scitranslmed.aat4271

Boulangé, C.L., Neves, A.L., Chilloux, J., Nicholson, J.K. & Dumas, M. (2016). Impact of the gut microbiota on inflammation, obesity, and metabolic disease. *Genome Medicine*, 8, 42. https://doi.org/10.1186/s13073-016-0303-2.

[dataset]Castro-Mejía, J.L., Khakimov, B., Krych, L., Bülow, J., Bechshøft, R.L., Højfeldt, G., … Nielsen, D. (2019). Physical fitness in community dwelling older adults is linked to dietary intake, gut microbiota and metabolomic signatures. European Nucletide Archive (ENA), accession number PRJEB33008.

Claesson, M.J., Jeffery, I.B., Conde, S., Power, S.E., O'Connor ,E.M., Cusack, S., … O'Toole, P.W. (2012). Gut microbiota composition correlates with diet and health in the elderly. *Nature,* 488, 178–184. https://doi.org/10.1038/nature11319

Dowd, K.P., Harrington, D.M. & Donnelly, A.E. (2012). Criterion and Concurrent Validity of the activPAL™ Professional Physical Activity Monitor in Adolescent Females. *PLoS One,* 7, e47633. https://doi.org/10.1371/journal.pone.0047633

Fei, N. & Zhao, L. (2013). An opportunistic pathogen isolated from the gut of an obese human causes obesity in germfree mice. *ISME Journa,* 7, 880–884. http://dx.doi.org/10.1038/ismej.2012.153.

Franceschi, C. & Campisi, J. (2014). Chronic Inflammation (Inflammaging) and Its Potential Contribution to Age-Associated Diseases. *Journals of Gerontology Series A,* 69, S4–9. https://doi.org/10.1093/gerona/glu057

Fujisaka, S., Avila-Pacheco, J., Soto, M., Kostic, A., Dreyfuss, J.M., Pan, H., … Kahn, C.R. (2018). Diet, Genetics, and the Gut Microbiome Drive Dynamic Changes in Plasma Metabolites. *Cell Reports,* 22, 3072–3086. https://doi.org/10.1016/j.celrep.2018.02.060

538 Gardner, D.G. & Shoback, D. (2011). *Greenspan's Basic & Clinical Endocrinology* (9th Ed).
539 McGraw-Hill Medical.

540 Giezenaar, C., Chapman, I., Luscombe-Marsh, N., Feinle-Bisset, C., Horowitz, M. & Soenen,
541 S. (2016). Ageing is associated with decreases in appetite and energy intake— A meta-
542 analysis in healthy adults. *Nutrients,* 8, 28. https://doi.org/10.3390/nu8010028

543 González, I., Cao, K.L., Davis, MJ & Déjean, S (2012). Visualising associations between
544 paired "omics" data sets. *BioData Minining,* 5, 19. https://doi.org/10.1186/1756-0381-5-
545 19

546 Goodrich, J.K., Waters, J.L., Poole, A.C., Sutter, J.L., Koren, O., Blekhman, R., … Ley, R.E.
547 (2014). Human genetics shape the gut microbiome. *Cell,* 159, 789–799.
548 https://doi.org/10.1016/j.cell.2014.09.053

549 Gopinath, B., Russell, J., Kifley, A., Flood, V.M. & Mitchell, P. (2016). Adherence to Dietary
550 Guidelines and Successful Aging Over 10 Years. *Journals of Gerontology Series A,* 71,
551 349–355. https://doi.org/10.1093/gerona/glv189

552 Guralnik, J.M., Simonsick, E.M., Ferrucci, L., Glynn, R.J., Berkman ,L.F., Blazer, D.G., …
553 Wallace, R.B. (1994). A short physical performance battery assessing lower extremity
554 function: Association with self-reported disability and prediction of mortality and
555 nursing home admission. *Journal of Gerontology,* 49, 85–94.

556 Hoarau, G., Mukherjee, P.K., Gower-Rousseau, C., Hager, C., Chandra, J., Retuerto, M.A.,
557 … Ghannoum, M.A (2016). Bacteriome and Mycobiome Interactions Underscore
558 Microbial Dysbiosis in Familial Crohn's Disease. *MBio,* 7, pii: e01250-16.
559 https://doi.org/10.1128/mBio.01250-16

560 Holm, L., Jespersen, A.P., Nielsen, D.S., Frøst, M.B., Reitelseder, S., Jensen, T., …
561 Damsholt, T .(2014). Hurrah for the increasing longevity: feasible strategies to
562 counteract age-related loss of skeletal muscle mass. *Scandinavian Journal of. Medicine*
563 *and Science in Sports,* 25, 1–2. https://doi.org/10.1111/sms.12415

564 Hothorn, T., Hornik, K., Strobl, C. & Zeileis, A. (2019). Party: a laboratory for recursive
565 partytioning. *R Packag.*

566 Husson, F., Josse, J., Lê, S & Mazet, J. (2019). Package 'FactoMineR.' *R Packag.*

567 Jones, C.J., Rikli, R.E. & Beam, W.C. (1999). A 30-s chair-stand test as a measure of lower
568 body strength in community-residing older adults. *Research Quaterly for Exercise and*
569 *Sport,* 70, 113–119.

570 Kashiwagi, I., Morita, R., Schichita, T., Komai, K., Saeki, K & Matsumoto, M (2015). Smad2
571 and Smad3 Inversely Regulate TGF- b Autoinduction in Clostridium butyricum -
572 Activated Dendritic Cells. *Immunity* 43., 65–79.
573 https://doi.org/10.1016/j.immuni.2015.06.010

574 Khan, MT., Nieuwdorp, M & Bäckhed, F (2014). Microbial modulation of insulin sensitivity.
575 *Cell Metabolism,* 20, 753–760. https://doi.org/10.1016/j.cmet.2014.07.006

576 Khan, SS., Singer, BD & Vaughan, DE (2017). Molecular and physiological manifestations
577 and measurement of aging in humans. *Aging Cell,* 16, 624–633.
578 https://doi.org/10.1111/acel.12601

579 Kong, F., Hua Y., Zeng, B., Ning, R., Li, Y & Zhao, J (2016). Gut microbiota signatures of

longevity. *Current Biology,* 26, R832–R833. https://doi.org/10.1016/j.cub.2016.08.015

Liaw, A (2018). Breiman and Cutler's Random Forests for Classification andRegression. *R Packag.*

Lustgarten, MS., Price, LL., Chalé, A & Fielding, RA (2014). Metabolites related to gut bacterial metabolism., peroxisome proliferator-activated receptor-alpha activation., and insulin sensitivity are associated with physical function in functionally-limited older adults. *Aging Cell,* 13, 918–925. https://doi.org/10.1111/acel.12251

Morrison, DJ., Preston, T., Morrison, DJ & Preston, T (2016). Formation of short chain fatty acids by the gut microbiota and their impact on human metabolism. *Gut Microbes,* 7, 189–200. https://doi.org/10.1080/19490976.2015.1134082

NIH (2011). *Global Health and Aging*. National Institute of Health.

Nordic Council of Ministers (2012). Nordic Nutrition Recommendations 2012. *Norden*.

Oksanen, J., Blanchet, F.G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., …Wagner, H (2019). vegan: Community Ecology Package. *R Packag.*

Peters, A., Krumbholz, P., Jäger, E., Heintz-Buschart, A., Çakir, MV., Rothemund, S., … Stäubert, C. (2019). Metabolites of lactic acid bacteria present in fermented foods are highly potent agonists of human hydroxycarboxylic acid receptor 3. *PLOS Genetics,* 15, e1008145. https://doi.org/10.1371/journal.pgen.1008145

Schacht, S.R., Lind, M.V., Bechshøft, R.L., Højfeldt, G., Reitelseder, S., Jensen, T., … Tetens, I. (2019). Investigating Risk of Suboptimal Macro and Micronutrient Intake and Their Determinants in Older Danish Adults with Specific Focus on Protein Intake—A Cross-Sectional Study. *Nutrients*, 11, pii: E795. https://doi.org/10.3390/nu11040795

Segerström, Å.B., Elgzyri, T., Eriksson, K.F., Groop, L., Thorsson, O & Wollmer, P (2011). Exercise capacity in relation to body fat distribution and muscle fibre distribution in elderly male subjects with impaired glucose tolerance., type 2 diabetes and matched controls. *Diabetes Research and Clinical Practice,* 94, 57–63. https://doi.org/10.1016/j.diabres.2011.05.022

Straub, L.G. & Scherer, P.E. (2019). Metabolic Messengers: adiponectin. *Nature Metaboism,* 1, 334–339. https://doi.org/10.1038/s42255-019-0041-z

Su, X., Yan ,H., Huang, Y., Yun, H., Zeng, B., Wang, E., … Yang, R. (2015). Expression of FABP4., adipsin and adiponectin in Paneth cells is modulated by gut Lactobacillus. *Scientific Reports,* 5, 18588. https://doi.org/10.1038/srep18588

Sumner, L.W., Amberg, A., Barrett, D., Beale, MH., Beger, R., Daykin ,C.A., … Viant, M.R. (2007). Proposed minimum reporting standards for chemical analysis. *Metabolomics*, 3, 211–221. https://doi.org/10.1007/s11306-007-0082-2

Thaiss, C.A., Zmora, N., Levy, M. & Elinav, E. (2016). The microbiome and innate immunity. *Nature,* 535, 65–74. https://doi.org/10.1038/nature18847

Tudor-Locke, C., Craig, C.L., Aoyagi, Y., Bell, R.C., Croteau, K.A., De Bourdeaudhuij, I., … Blair ,S.N. (2011). How many steps/day are enough? For older adults and special populations. *International Journal of Behavioral Nutrition and Physical Activity,* 8, 80. https://doi.org/10.1186/1479-5868-8-80

WHO (2000). Obesity: preventing and managing the global epidemic. World Health Organization

Xue, Q.L. (2011). The Frailty Syndrome: Definition and Natural History. *Clinics in Geriatric Medicine,* 27, 1–15. https://doi.org/10.1016/j.cger.2010.08.009

Yang, Y.J., Kim, M.K., Hwang, S.H., Ahn, Y., Shim, J.E. & Kim, D.H. (2010). Relative validities of 3-day food records and the food frequency questionnaire. *Nutrition Research in Practice,* 4, 142–148. https://doi.org/10.4162/nrp.2010.4.2.142

Zeevi, D., Korem, T., Zmora, N., Israeli, D., Rothschild, D., Weinberger, A., … Segal, E. (2015). Personalized Nutrition by Prediction of Glycemic Responses. *Cell,* 163, 1079–1095. https://doi.org/10.1016/j.cell.2015.11.001

SUPPORTING INFORMATION

Additional supporting information can be found in supplementary files.

- Figure S1. Data overview and dietary intake
- Figure S2. GM overview, cumulative- and core-species
- Figure S3. rCC analysis between GM and lifestyle components
- Figure S4. Metabolome correspondence and correlation
- Table S1. Subjects Stratification
- Table S2. Dietary evaluation
- Table S3. Proinsulin and HbA1c levels
- Supplementary Methods

652

653 # TABLES

654 Table 1. Description of the study participants

Number of Participants (n)	207
Sex	
Men:Women	109:98
Age (y) Mean ± SD	70.2 ± 3.9
BMI (kg·m^2) Mean ± SD	25.7 ± 3.8
BMI < 25	105
BMI ≥ 25 < 30	75
BMI ≥ 30	27
HbA1c (mmol·mol^{-1})	
< 39 mmol·mol^{-1} (<5.7 ABG - mmol·l^{-1}) [a]	167
39-46 mmol·mol^{-1} (5.7-6.4 ABG - mmol·l^{-1})	40

655

656 [a] HbA_{1c} values above 47 mmol·mol^{-1} (6.5 mmol·l^{-1} Average Blood Glucose - ABG) is a

657 criterion for diagnosis of T2D (Gardner & Shoback 2011).

658

659

660 Table 2. Within sex summary of ABP measurements used for stratification of phenotypes

661 (HF/P: high-fitness phenotypes, LF: low-fitness phenotypes).

Women

Functional-Parameter	HF/P	LF/P	*p*-value[a]	Refer. range	Ref. age
30s Chair-stand test	20.6 ± 5.0	15.7 ± 3.1	< 0.001	10–16[b]	65-74y [b]
BMI	22.4 ± 2.1	28.9 ± 3.3	< 0.001		
LSF%	35.2 ± 4.0	42.7 ± 4.6	< 0.001		

Men

Functional-Parameter	HF/P	LF/P	*p*-value[a]	Refer. range	Ref. age
Chair-rise test	22.9 ± 4.4	18.3 ± 3.9	< 0.001	12–18[b]	65-74y [b]
BMI	24.0 ± 2.2	28.3 ± 3.1	< 0.001		
LSF%	20.3 ± 3.4	27.0 ± 3.5	< 0.001		

662

663 [a] Comparison between phenotypes was performed by two-tailed Student's *t*-test.

664 [b] ref: (Jones et al., 1999; Guralnik et al., 1994)

665

666 ## FIGURE LEGENDS

667 **Figure 1. Stratification of fitness phenotypes**

668 (a) Stratification of subjects (n = 207) by hierarchical clustering analysis of principal
669 components analysis (HCA-PCA). Stratification data matrix: [obj x vars] = [207 x 3]. HCA-
670 PCA was performed within sexes and based on ABP measurements. HF/P: high-fitness (n =
671 116) and LF/P: low-fitness phenotypes (n = 91).

672 (b) ABP measurements distribution among phenotypes and sexes.

673 (c) 4-day activity monitoring displaying hours standing and steps on daily basis for both
674 phenotypes. 4-day activity data matrix: [obj x vars] = [196 x 2]

675

676

677

678

679

680

681

682

683

684

685

686

687

688

689

690

Figure 2. Dietary intake and distribution

(a) Total energy consumption per kg-body-weight per day (Cal·kg-body-weight^{-1}·day^{-1})

(b) Distribution of Calories proportionally obtained from macronutrients intake in HF and LF phenotypes.

(c) Intake of carbohydrates by quality and saturated free fatty acids (g·kg-body-weight^{-1}·day^{-1}).

(d) Pearson correlation between dietary fiber (g·kg-body-weight^{-1}·day^{-1}) and BMI depicted according to phenotypes category.

(e) Proportion of subjects complying with recommended carbohydrates distribution ranges. The gray areas correspond to non-recommended ranges as suggested by the Nordic Nutrition Recommendations.

(f) Proportion of subjects complying with recommended distribution ranges of dietary fiber according to the Nordic Nutrition Recommendations.

Dietary data matrix: [obj x vars] = [181 x 11]

718

719

720

721 **Figure 3. Dietary intake and fitness phenotypes is linked with species-level GM patterns**

722 (a) Gut microbiota (GM) composition determined through Correspondence Analysis of 16S
723 rRNA gene (V3-region) amplicons (summarized zOTUs at species level) determined in the
724 stool samples of the study participants.
725 (b) Correspondence Analysis revealed compositional GM differences between fitness
726 phenotypes. (c) Constrained Correspondence Analysis (CCA) displays discrimination of
727 phenotypes based on permutational test (p = 0.03, explained variance = 3.2%).
728 (d) Correspondence Analysis of GM composition depicting gradients of total energy
729 consumption (Cal·kg-body-weight^{-1}·day^{-1}), intake of (e) starch (g·kg-body-weight^{-1}·day^{-1}) and
730 (f) dietary fiber (g·kg-body-weight^{-1}·day^{-1}), (g) steps per day, and (h) BMI.
731 (i) rCC analysis depicting the relationship between gradients of energy consumption, starch
732 and dietary fiber intake, steps per day and BMI, and variations in the abundance of GM
733 members. Heatmap displays the correlation of 161 species with a minimum correlation
734 coefficient of |0.2|r from 1st to 3rd components. Species are depicted based on family-level
735 phylogeny. Supplementary Figure 3 displays taxonomy at species level, as well as
736 correlations per canonical axis and explained variance between GM composition and lifestyle
737 covariates derived from rCC analysis.
738 ANOSIM tests were performed on Bray-Curtis distances. GM data matrix: [obj x vars] = [184
739 x 874]

740

741

742

743

744

Figure 4. Profiling of host metabolome in relation to dietary intake

(a) Correspondence Analysis on combined fecal-, plasma-metabolomes and clinical biomarkers of the study participants. Significant differences due to sex were determined with constrained correspondence analysis (CCA). Inset shows a partial Correspondence Analysis after conditioning for the cofounding effect of sex.

(b) Correspondence Analysis discriminates compositional differences in metabolomic profiles between fitness phenotypes.

(c) Correspondence Analysis of metabolites in relation to total energy consumption (Cal·kg-body-weight^{-1}·day^{-1}), intake of (d) dietary fiber (g·kg-body-weight^{-1}·day^{-1}), (e) starch (g·kg-body-weight^{-1}·day^{-1}) and (f) simple sugars (g·kg-body-weight^{-1}·day^{-1}), (g) steps per day, (h) hours standing, and (i) BMI.

(j) rCC analysis showing the relationship between gradients of energy consumption, dietary fiber, starch and simple sugar intake, steps per day, hours standing and BMI, with variations in metabolome composition. Heatmap displays the correlation of 34 clinical/metabolome variables with a minimum correlation coefficient of $|0.2|r$ from 1st to 4th components. Supplementary Figure 4 shows correlations per canonical axis as well as explained variance between metabolome composition and lifestyle covariates derived from rCC analysis.

(k) Significantly (t-test, $p = 0.02$) different relative distributions in enterolactone determined in fecal samples of HF and LF phenotypes

(l-m) Range of fecal SCFAs and O/B-CFAs concentrations sorted according to fitness phenotype.

Metabolome data matrix: [obj x vars] = [184 x 335]

772

773

774 **Figure 5. Signatures discriminating physical phenotypes**

775 (a) Heatmap displaying mean centered normalized abundance of 56 features selected using
776 Random Forest towards discrimination of phenotypes and (b) their importance as determined
777 on the basis of Mean Decrease in Accuracy.
778 (c) Multidimensional scaling plot discriminates subjects' phenotype based on the selected
779 features.
780 (d) ROC curves and out-of-bag error rate (OOB) for Random Forest classifier based on the
781 selected variables, for combined datasets (all selected features), GM and metabolome, dietary
782 intake and physical activity
783 (e) Captured variance for fitness variables (BMI, chair-stand and LSF%) as a function of
784 selected features through redundancy analysis (RDA). Individual Explained Variance displays
785 the size effect of a given dataset, CE-Variance represents the cumulative explained variance
786 and CE-variance | physical-activity shows the accumulative explained variance conditioned
787 by physical activity. Pie charts summarize the total proportion of explained variance before
788 and after conditioning for physical activity.
789 Data matrix: [obj x vars] = [181 x 56]

Figure 1

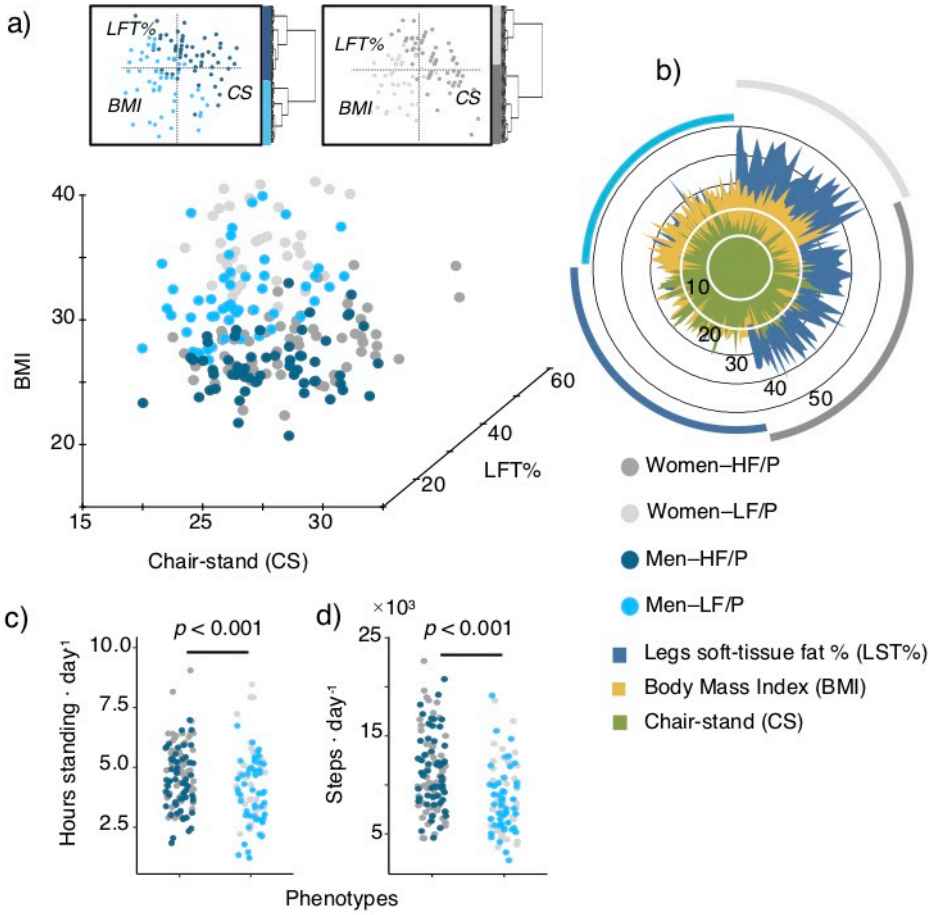


Figure 2

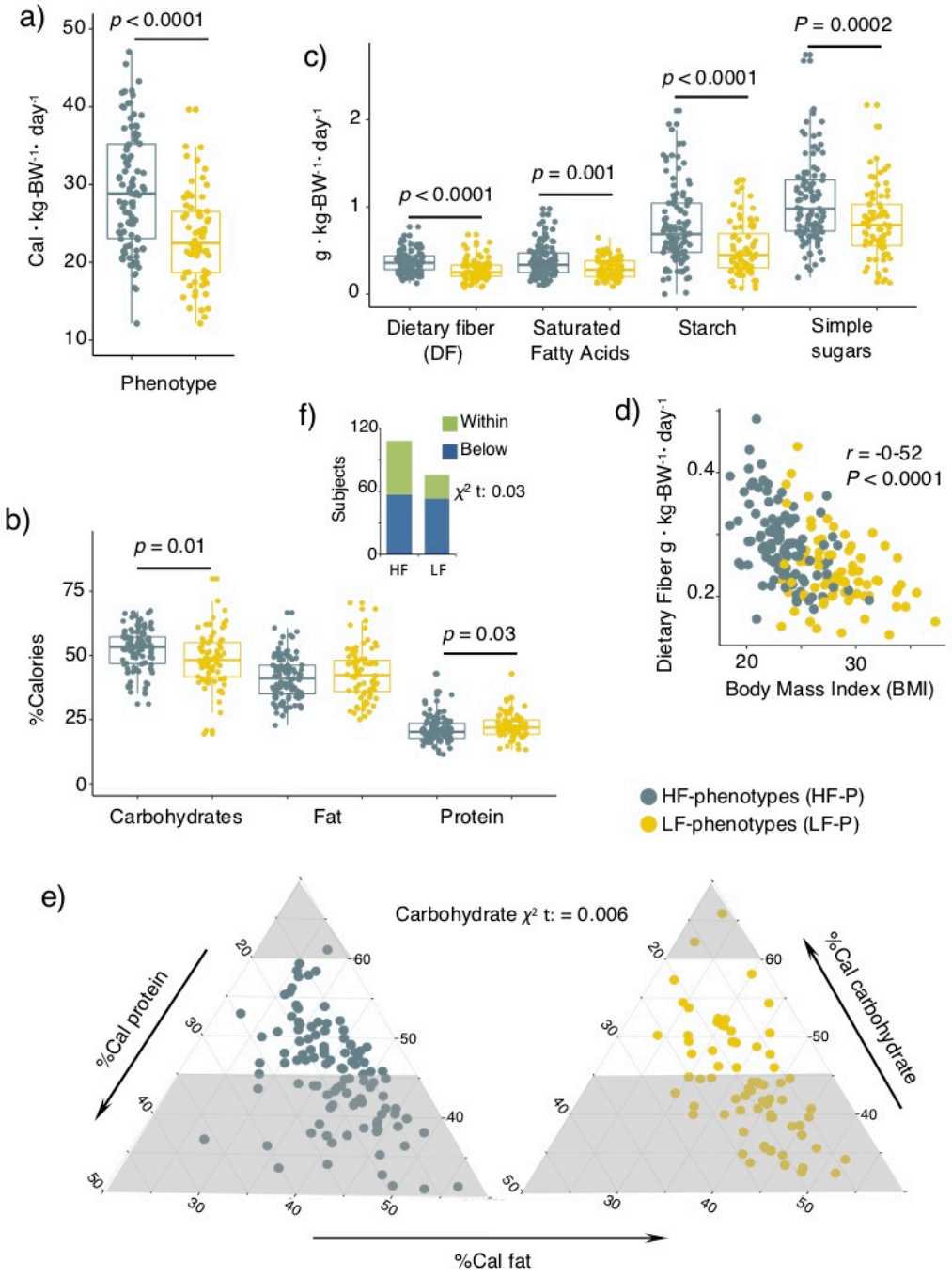


Figure 3

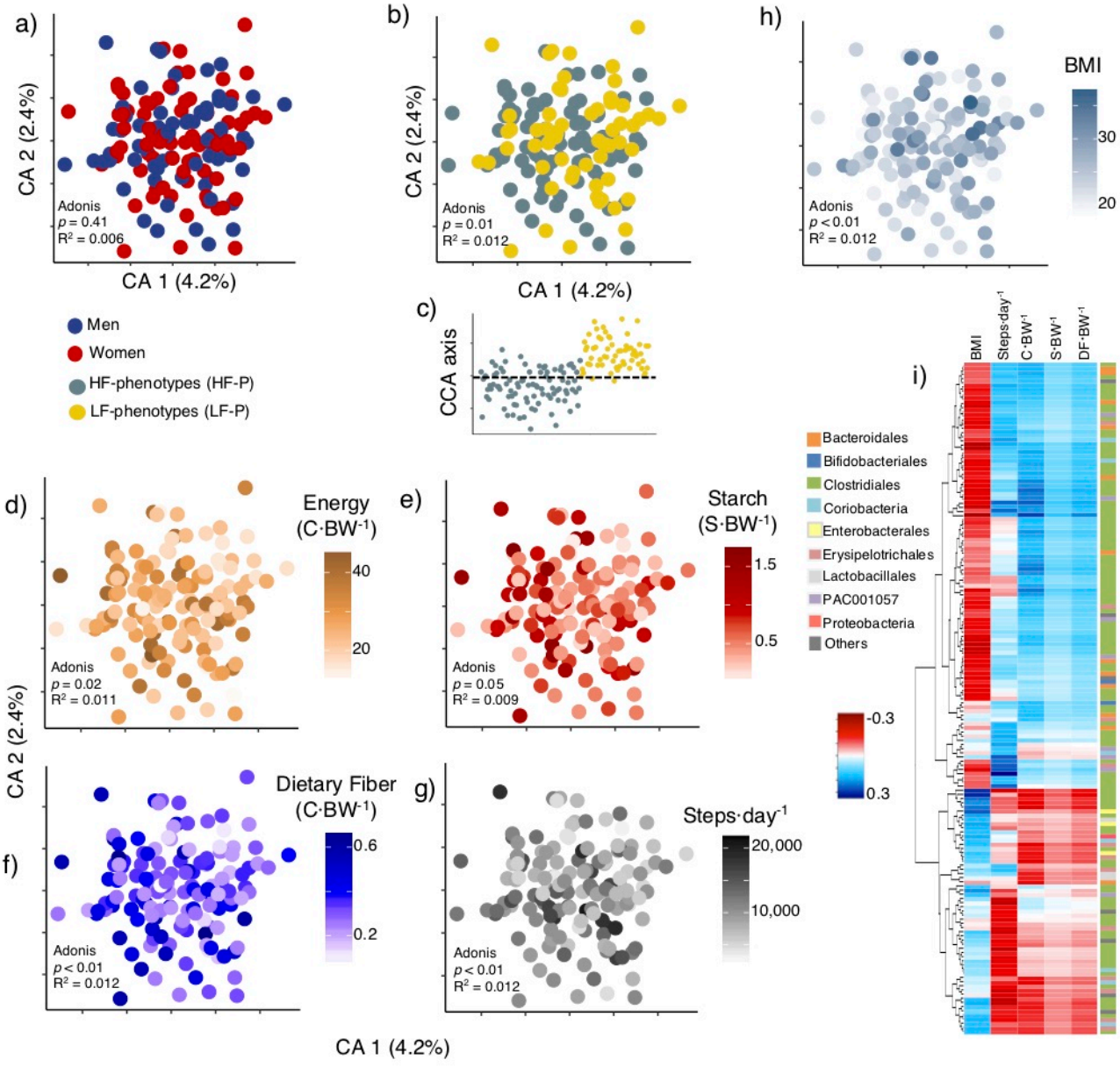


Figure 4

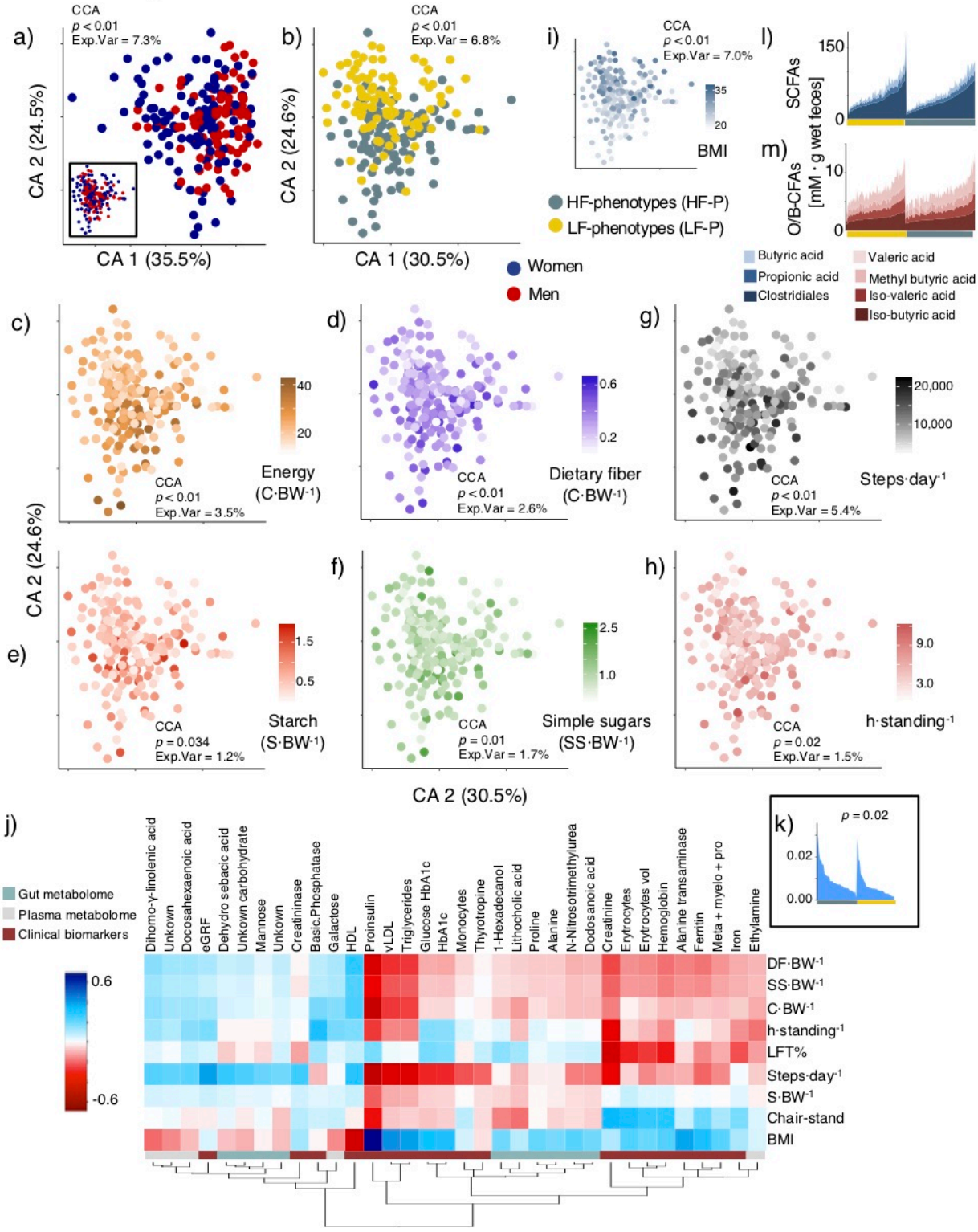
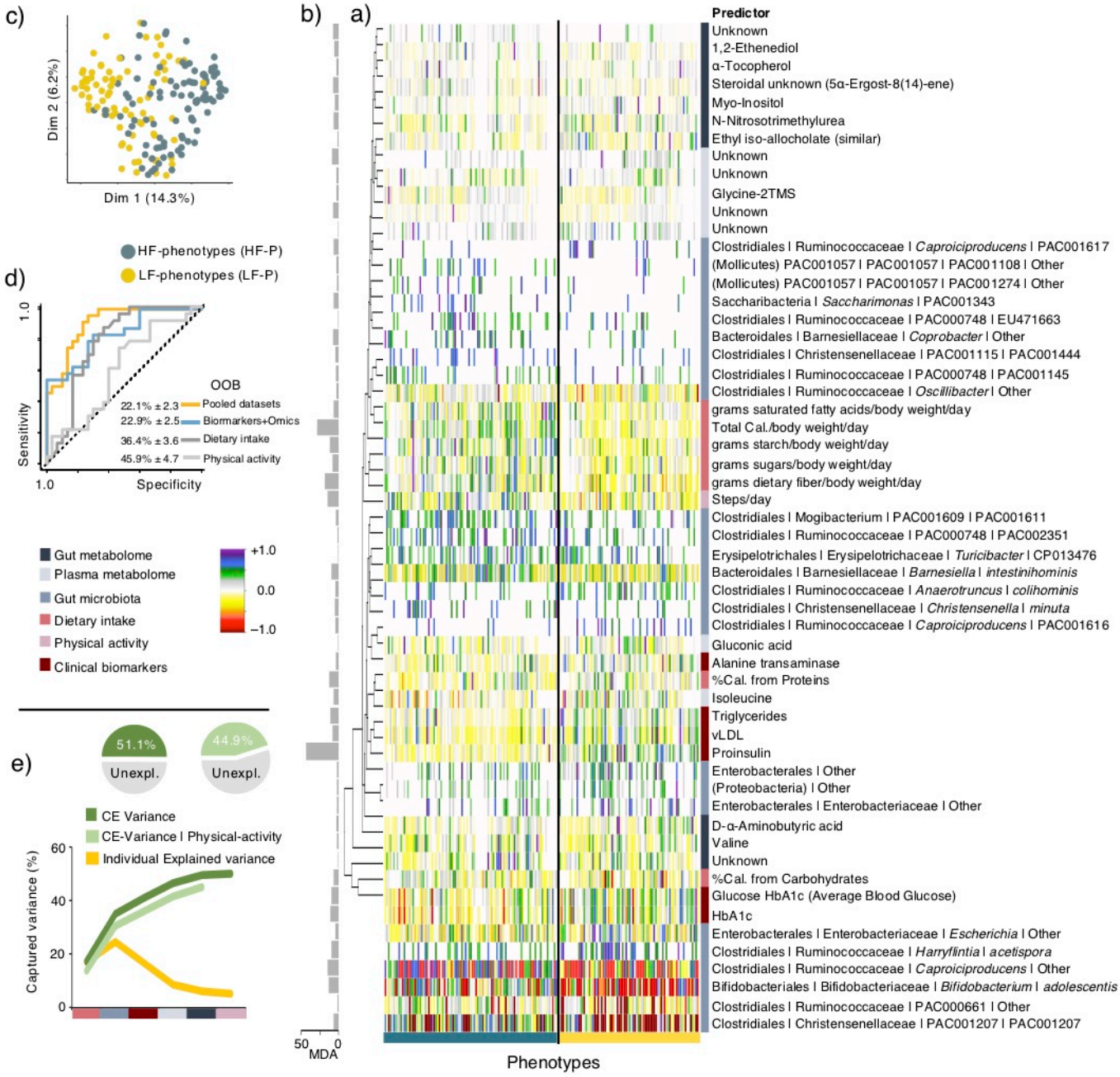
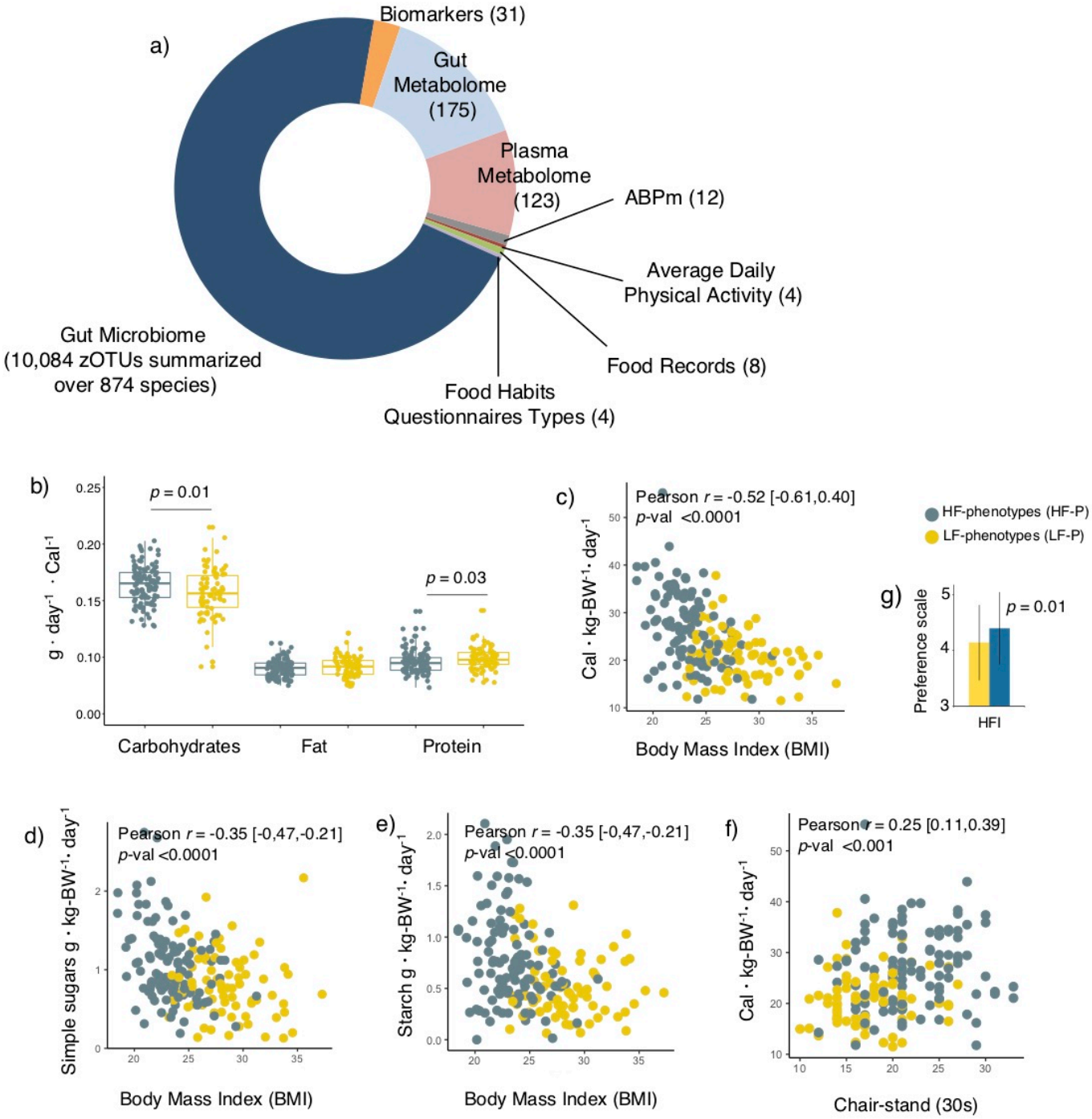


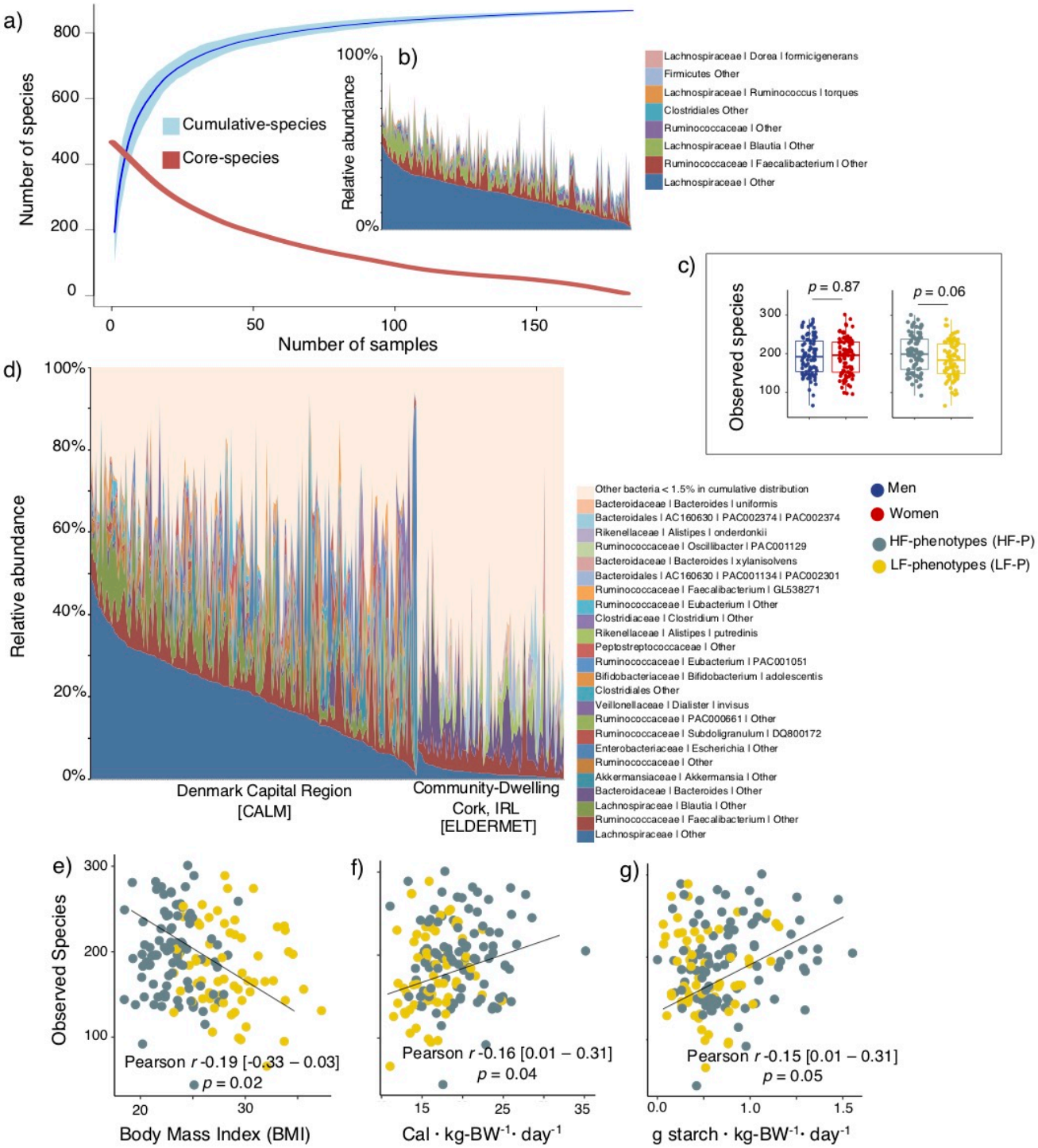
Figure 5





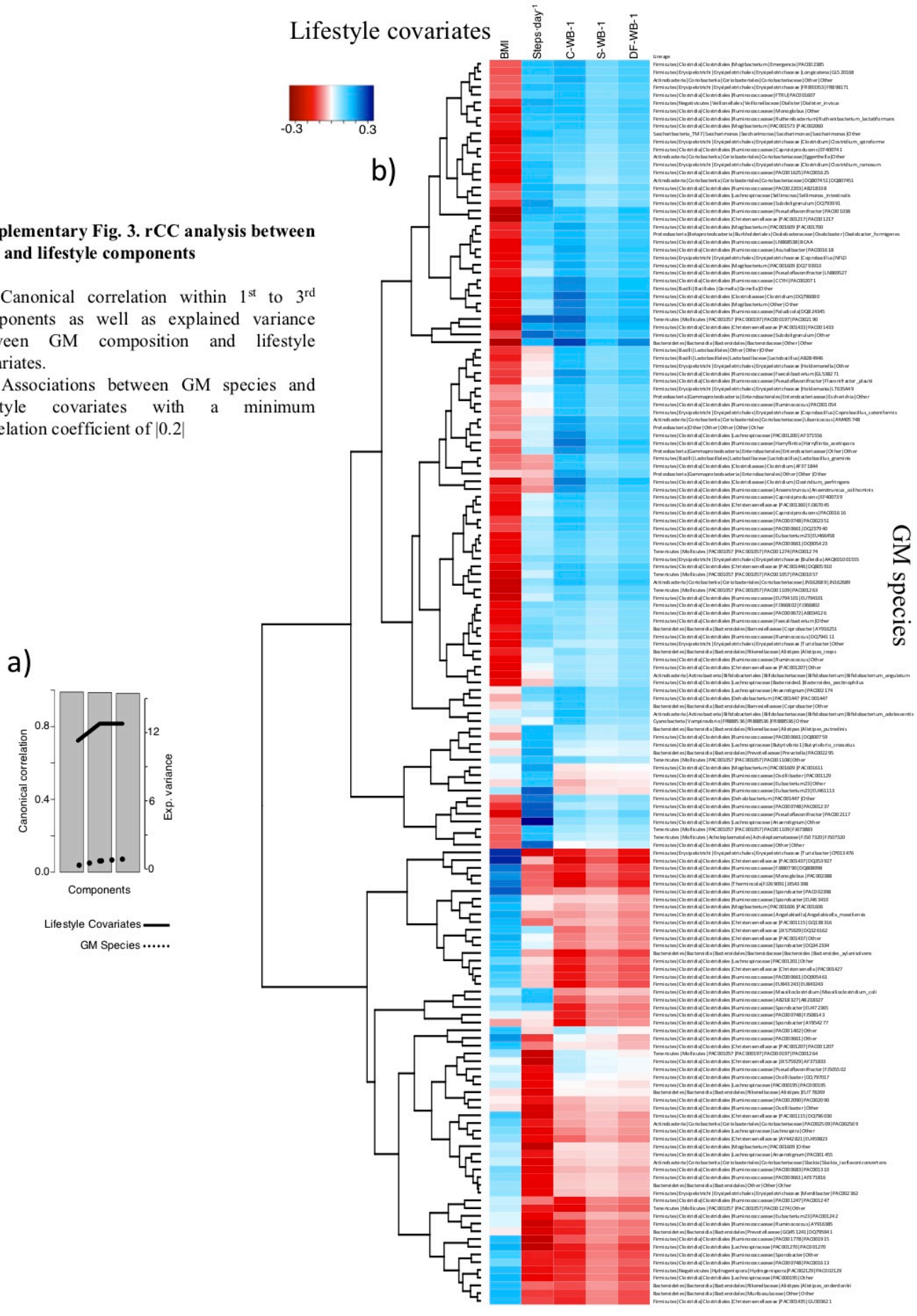
Supplementary Fig. 1. Subjects characterization and dietary intake

(a) Ring chart displays the proportion of variables used for every category at which individuals were characterized
(b) Distribution of daily macronutrient intake (g day^{-1}) normalized by the total energy intake (Cal)
(c) Correlation between energy intake (Cal · kg-BW^{-1}· day^{-1}) *vs* BMI
(d) Correlation between starch intake (g · kg-BW^{-1}· day^{-1}) *vs* BMI
(e) Correlation between simple sugars intake (g · kg-BW^{-1}· day^{-1}) *vs* BMI
(f) Correlation between energy intake (Cal · kg-BW^{-1}· day^{-1}) *vs* Chair-stand test
(g) Degree of agreement for food-choices questionnaires: healthy food in an important element of everyday life (HFI). This was evaluated on a scale of 1 – 5. 1: is not important, 5: is very important,



Supplementary Fig. 2. GM overview, cumulative- and core-species

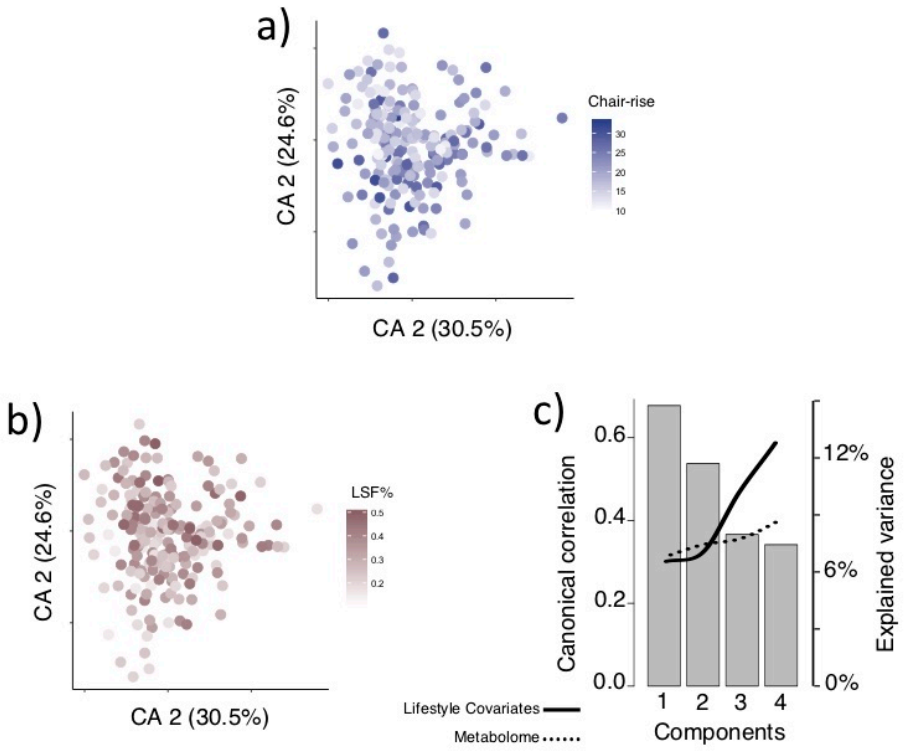
(a) Cumulative- and core species comprised in the CALM study with increasing number of subjects.
(b) Relative abundance of core-species (phylotypes summarized to species level).
(c) Alpha diversity (of summarized zOTUs at species level) between sexes ($p = 0.42$) and phenotypes ($p = 0.04$) determined by Monte Carlo permutation (100) test.
(d) Distribution of species across the subjects of the CALM intervention (16S rRNA v3 region) and the community-dwellers of the ELDERLMET study (16S rRNA v4 region, 454 pyro-sequencing). The publically available sequencing data from the ELDERMENT study were retrieved and analyzed using the parameters described in methods.
(e) Correlation between observed species *vs* BMI
(f) Correlation between observed species *vs* Energy intake (Cal · kg-BW^{-1}· day^{-1})
(g) Correlation between observed species *vs* Simple sugars intake (g · kg-BW^{-1}· day^{-1})



Supplementary Fig. 3. rCC analysis between GM and lifestyle components

(a) Canonical correlation within 1st to 3rd components as well as explained variance between GM composition and lifestyle covariates.

(b) Associations between GM species and lifestyle covariates with a minimum correlation coefficient of |0.2|



Supplementary Fig. 4. Metabolome correspondence and correlation

(a) Correspondence Analysis of metabolome in relation to chair-stand test and (b) LSF%
(c) Canonical correlation within 1st and 4th components as well as explained variance between metabolome profiling and lifestyle covariates.

Fecal DNA extraction, 16S rRNA-gene amplicon sequencing

Fecal samples were thawed at 4°C, re-suspended in ultrapure water (1:2 feces/water) and homogenized in filter bags for 1 min at high speed (Lab Seward, BA7021). 1.5 ml of the fecal slurry was centrifuged at 13,000×g for 10 min at room temperature and ~200 mg of the fecal pellet was used for DNA extraction using the PowerSoil® DNA Isolation Kit (MOBIO Laboratories, Carlsbad, CA, USA), basically following the instructions of the manufacturer, but with minor modifications to increase lysis of bacterial cells: prior DNA extraction, samples were placed into the PowerBead tubes and heat treated at 65°C for 10 min and then at 95°C for 10 min. Subsequently, solution C1 was added and bead-beating performed in FastPrep (MP Biomedicals, Santa Ana, CA, USA) using 3 cycles of 15 s each, at a speed of 6.5 m s^{-1}. The remaining DNA extraction procedure followed the manufacturer's instructions. Gut prokaryotic composition was determined by NexSeq 500 based 16S rRNA gene-amplicon sequencing of the V3 region amplified using primers designed with adapters for the Nextera Index Kit® (Illumina, CA, USA): NXt_338_F: 5'- TCG TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG ACW CCT ACG GGW GGC AGC AG -3' and NXt_518_R: 5'- GTC TCG TGG GCT CGG AGA TGT GTA TAA GAG ACA GAT TAC CGC GGC TGC TGG -3'. Amplification profile (1^{st} PCR), barcoding (2^{nd} PCR), amplicon library purification and sequencing were performed as previously described (Pyndt Jørgensen et al. 2014).

Analysis of high-throughput amplicon sequencing

The raw dataset containing pair-ended reads with corresponding quality scores were merged and trimmed using the following settings, -fastq_minovlen 100, -fastq_maxee 2.0, -fastq_truncal 4, -fastq_minlen 130. Finding unique reads and deconvoluting from chimeric reads and constructing *de-novo* zero-radius Operational Taxonomic Units (zOTU) was conducted using the UNOISE pipeline (Edgar 2018) coupled to the EZtaxon 16S rRNA gene collection as a reference database (Kim et al. 2012). Downstream analyses were based on a contingency table

rarefied to 17,000 sequences per sample and then normalized with cumulative sum scaling (CSS (Paulson et al. 2013)).

Metabolomics

Untargeted metabolomics of fecal slurries

1 ml fecal homogenate (as described above) was mixed with 1 ml of Sterile PBS (5.7 mM Na_2HPO_4, 24.3 mM NaH_2PO_4, 450 mM NaCl, pH 7.4), frozen in liquid nitrogen and freeze-dried overnight. Twenty mg of each freeze-dried sample were re-suspended in 1 ml of 99.98% methanol (containing 10 ppm palmitic-acid methyl ester and 10 ppm sorbitol as an internal standards), vortexed and centrifuged for 30 min at 12,000×g at 4°C. Fifty µl of the supernatant were then dried using a ScanVac (Labogene, Lynge, Denmark) at 1,000 rpm for 3 h at 40°C. Immediately after drying, samples were sealed with air tight magnetic lids into 2.0 ml GC-MS vials and derivatized in two steps using a Dual-Rail MultiPurpose Sampler (MPS) (Gerstel, Mülheim an der Ruhr, Germany), (i) addition of 10 µl of MEOX reagent (20 mg ml^{-1} Methoxiamine hydrochloride in dry pyridine) followed by agitation at 45°C for 90 min by mixing at 750 rpm, (ii) addition of 40 µl of TMS reagent, trimethylsilyl cyanide (TMSCN) (Khakimov et al. 2013) followed by agitation at 45°C for 45 min by mixing at 750 rpm. All steps involving sample derivatization and injection were automated using MPS, which was equipped with a sample agitation unit. Immediately after derivatization, 1 µl of the derivatized sample was injected into a cooled injection system (CIS4) (Gerstel, Mülheim an der Ruhr, Germany) port in splitless mode. The septum purge flow and purge flow to split vent at 2.5 min after injection were set to 25 and 15 ml min^{-1}, respectively. Initial temperature of the CIS4 port was 45°C, and heated at 12°C s^{-1} to 320°C (after 30 s of equilibrium time), where it was kept for 10 min. After heating, the CIS4 port was gradually cooled to 250°C at 5°C s^{-1}, and this temperature was kept constant during the run. The GC-TOF-MS setup was made combining an Agilent 7890B gas chromatograph (GC) (Agilent Technologies, California, USA) with a time-of-flight mass spectrometer, HT Pegasus TOF-MS, (LECO Corporation, Saint Joseph, USA).

GC separation was performed on a Zebron ZB 5% Phenyl 95% Dimethylpolysiloxane column (30 m with I.D. 250 µm and film thickness 0.25 µm) with a 5 m inactive guard column (Phenomenex, Torrance, USA). A hydrogen generator, Precision Hydrogen Trace 500 (Peak Scientific Instruments Ltd, Inchinnan, UK) was used to supply a carrier gas at a constant column flow rate of 1.0 ml min^{-1}. The initial temperature of the GC oven was set to 40°C and held for 2 min followed by heating at 10°C min^{-1} to 320°C and kept for an additional 6 min, making the total run time 36 min. Mass spectra was recorded in the range of 45–600 m/z with a scanning frequency of 10 scans sec^{-1}, and the MS detector and ion source was switched off during the first 6.3 min of solvent delay time. The transfer line and ion source temperature were set to 280°C and 250°C, respectively. The mass spectrometer was tuned according to manufacturer's recommendation using perfluorotributylamine (PFTBA). MPS and GC-TOF-MS were controlled using vendor software Maestro (Gerstel, Mülheim an der Ruhr, Germany) and ChromaTOF (LECO Corporation, Saint Joseph, USA), respectively. Samples were randomized prior to derivatization and GC–MS analysis. In order to monitor instrument performance, a blank sample containing only derivatization reagent, a control sample (a pooled sample), and an alkane mixture standard sample (all even C10-C40 alkanes at 50 mg L^{-1} in hexane) were injected after every 10 real samples.

The raw GC-TOF-MS data was processed using Statistical Compare toolbox of the ChromaTOF software (Version 4.50.8.0) with following settings; the raw data was used without smoothing prior to peak deconvolution, baseline offset was set to 0.8, expected averaged peak width was set to 1.5 sec, signal-to-noise was set to ≥10, peak areas were calculate using deconvoluted mass spectra (DT), common *m/z* ions of derivatization products were determined as 73, 75, and 147, deconvoluted mass spectra were also used for peak identification using LECO-Fiehn and NIST11 libraries. The library search was set to return top 10 hits with EI-MS match of >75% using normal-forward search and with a mass threshold of 20. Deconvoluted peaks were aligned across all samples using following settings; retention time shift allowance of <3 sec, EI-MS match of >95%, mass threshold of >25, and present in >90% of all pooled control samples.

Targeted analysis of SCFA and O/B-CFA in fecal slurries

Analysis of SCFA and O/B-CFA was performed on 0.5 ml of fecal homogenate mixed with 1 ml of 0.3M oxalic acid (containing 2 mM of 2 ethylbutyrate (Sigma-Aldrich) as the internal standard). Samples were vortexed for 1 min, centrifuged at 20°C for 20 min at 12,000×g, followed by filtration using a 0.45 µm centrifugal filter (Millipore UFC30HV00) and the obtained aliquot was used for GC-MS analysis. The GC-MS consisted of an Agilent 7890A GC and an Agilent 5973 series MSD. GC separation was performed on a Phenomenex Zebron ZB-WAXplus column (30 m × 250 µm × 0.25 µm). A sample volume of 1 µl was injected into a split/splitless inlet at 285°C using split mode at 2:1 split ratio. Septum purge flow and split flow were set to 13 ml min^{-1} and 2 ml min^{-1}, respectively. Hydrogen was used as carrier gas, at a constant flow rate of 1.0 ml min^{-1}. The GC oven program was as follows: initial temperature 100°C, equilibration time 1.0 min, heat up to 120°C at the rate of 10°C min^{-1}, hold for 5 min, then heat at the rate of 40°C min^{-1} until 230°C and hold for 2 min. Mass spectra were recorded in Selected Ion Monitoring (SIM) mode and m/z ions were detected at the dwell time of 50 msec: 41, 43, 45, 57, 60, 73, 74, 84. The detector was switched off during the 1 min of solvent delay time. The transfer line, ion source and quadrupole temperatures were set to 230, 230 and 150°C, respectively. The mass spectrometer was tuned according to manufacturer's recommendation using perfluorotributylamine (PFTBA). Dilution series of SCFA standards of acetic, propionic, butyric, isobutyric, 2-methyl isobutyric, valeric and isovaleric acid (Sigma-Aldrich) were prepared in concentrations of 1.000, 0.500, 0.250, 0.125, 0.060 and 0.030 mM for the construction of standard curves for quantification. Initial inspection of the GC-MS data was performed using MSD ChemStation software (Version E.02.02.1431, Agilent Technologies, Inc., Germany). Mass spectra of SCFA were compared against the NIST11 library (NIST, Maryland, USA). SCFA peak areas were integrated from SIM chromatograms using in-house Matlab (Version. R2015a, The MathWorks, Inc., Massachusetts, USA) scripts. Two SCFA, 2-methyl isobutyric acid and isovaleric acid, co-eluted at the retention time range of 4.22-4.45

min, thus peak areas were calculated by deconvoluting these peaks using *m/z* ions 74 for 2-methyl isobutyric acid and 60 for isovaleric acid.

Untargeted metabolomics of blood plasma

A mixture of 100 µl of plasma samples (thawed at room temperature) and 300 µl of MeOH:water (8:1, vol:vol and containing 10 ppm of sorbitol as internal standard) were vortexed (highest speed) for 1 min. Thereafter, samples were incubated at 4°C for 15 min and centrifuged at 16,000×g at 4°C for 10 min. Supernatants were passed through a 0.45 µm centrifugal filter (Millipore UFC30HV00) and 80 µl aliquots were dried into 200 µl glass inserts using a ScanVac (Labogene, Lynge, Denmark) at 40°C for 3 h at 1,000 rpm. Immediately after drying samples were sealed with air tight magnetic lids into 2.0 ml GC-MS vials and derivatized in two steps using MPS, (i) addition of 10 µl of MEOX reagent (20 mg ml-1 Methoxiamine hydrochloride in dry pyridine) followed by agitation at 65°C for 60 min by mixing at 750 rpm, (ii) addition of 30 µl of TMS reagent (TMSCN) followed by agitation at 65°C for 2 h by mixing at 750 rpm. Immediately after derivatization, 1 µl of the derivatized sample was injected into the GC-TOF-MS as described for the fecal metabolomics. Sample injection, oven and mass spectrometer parameters were similar to those for the fecal metabolomics with few modifications. The initial temperature of the GC oven was set to 40°C and held for 2 min followed by heating at 12 °C min^{-1} to 260°C, and with a rate of 30°C min^{-1} to 320°C and kept for an additional 5 min, making the total run time 27.33 min. Mass spectra was recorded in the range of 45–500 m/z with a scanning frequency of 8 scans sec^{-1}, and the MS detector and ion source was switched off during the first 8.3 min of solvent delay time. The transfer line and ion source temperature were set to 290°C and 250°C, respectively. In order to monitor instrument performance, a blank sample containing only derivatization reagent, a control sample (a pooled sample), and an alkane mixture standard sample (all even C10-C40 alkanes at 50 mg L^{-1} in hexane) were injected after every 10 real samples. The raw GC-TOF-MS data was processed as described above for untargeted fecal metabolomics.

References

Edgar, R.C. (2018) Updating the 97% identity threshold for 16S ribosomal RNA OTUs A. *Bioinformatics* 34, 2371–2375. https://doi.org/10.1093/bioinformatics/bty113

Khakimov, B., Motawia ,M.S., Bak, S. & Engelsen, S.B. (2013) The use of trimethylsilyl cyanide derivatization for robust and broad-spectrum high-throughput gas chromatography-mass spectrometry based metabolomics. *Analytical and Bioanalytical Chemistry,* 405, 9193–9205. https://doi.org/10.1007/s00216-013-7341-z)

Kim, O.S., Cho, Y.J., Lee, K., Yoon, S.H., Kim, M., Na, H., … Chun, J. (2012) Introducing EzTaxon-e: a prokaryotic 16S rRNA gene sequence database with phylotypes that represent uncultured species. *International Journal of Systematic and Evolutionary Microbiology,* 62, 716–721. https://doi.org/10.1099/ijs.0.038075-0

Paulson, J.N., Stine, O.C., Bravo, H.C. & Pop, M. (2013) Differential abundance analysis for microbial marker-gene surveys. *Nature Methods* 10, 1200–1202. https://doi.org/10.1038/nmeth.2658

Pyndt Jørgensen, B., Hansen, J.T., Krych, L., Larsen, C., Klein, A.B., Nielsen, D.S., J… Sørensen, D.B. (2014) A Possible Link between Food and Mood: Dietary Impact on Gut Microbiota and Behavior in BALB/c Mice. *PLoS One* 9, e103398. https://doi.org/10.1371/journal.pone.0103398