



Variation of Carbohydrate-Active Enzyme Patterns in the Gut Microbiota of Italian Healthy Subjects and Type 2 Diabetes Patients

Matteo Soverini, Silvia Turroni, Elena Biagi, Sara Quercia, Patrizia Brigidi, Marco Candela and Simone Rampelli*

Unit of Microbial Ecology of Health, Department of Pharmacy and Biotechnology, University of Bologna, Bologna, Italy

OPEN ACCESS

Edited by:

Gayane Manukyan,
Institute of Molecular Biology
(NAS RA), Armenia

Reviewed by:

Maryam Dadar,
Razi Vaccine and Serum Research
Institute, Iran
Mario M. D'Elia,
University of Florence, Italy

*Correspondence:

Simone Rampelli
simone.rampelli@unibo.it

Specialty section:

This article was submitted to
Microbial Immunology,
a section of the journal
Frontiers in Microbiology

Received: 31 July 2017

Accepted: 11 October 2017

Published: 24 October 2017

Citation:

Soverini M, Turroni S, Biagi E,
Quercia S, Brigidi P, Candela M and
Rampelli S (2017) Variation
of Carbohydrate-Active Enzyme
Patterns in the Gut Microbiota
of Italian Healthy Subjects and Type 2
Diabetes Patients.
Front. Microbiol. 8:2079.
doi: 10.3389/fmicb.2017.02079

The human gut microbiota (GM) has been associated, to date, with various complex functions, essentials for the host health. Among these, it is certainly worth noting the degradation of the so-called microbiota-accessible carbohydrates (MACs), which the GM breaks down through specific enzymes, referred to as carbohydrate-active enzymes (CAZymes). This degradation constitutes the first step in the production of short-chain fatty acids (SCFAs), key microbial small molecules having multiple health-promoting effects for the host organism. The decline in MAC dietary intake in urban Western populations forced the shrinkage of CAZyme repertoire in the GM, as shown by the literature comparing the microbiome layout between Western urban citizens and traditional rural populations. Even if this reduction in GM functional complexity has been associated with the onset of the so-called “diseases of civilization,” only few information regarding the CAZyme variation within Western populations has been provided to date, and its connections with diet and health are still unexplored. In this scenario, here we explore the GM-encoded CAZyme repertoire across two Italian adult cohorts, including healthy lean subjects consuming a Mediterranean diet and obese patients affected by type 2 diabetes, consuming a high-fat diet. In order to impute the CAZyme panel, a pipeline consisting of publicly available software – QIIME, FragGeneScan and HMMER – was specifically implemented. Our study highlighted the existence of robust clusters of bacterial species sharing a common MAC degradation profile in the Italian GM, allowing the stratification of the individual GM into different steady states according to the carbohydrate degradation profile, with possible connections with diet and health.

Keywords: gut microbiota, microbiota-accessible carbohydrates, CAZymes, co-abundance groups, type 2 diabetes

INTRODUCTION

The large ensemble of bacteria that stably inhabit the distal part of our gastrointestinal tract, namely the gut microbiota (GM), is of vital importance for many physiological functions of our organism, exerting a key role in several biological processes, such as nutrition, immune function, and the regulation of the central nervous system (Neish, 2009; Hooper et al., 2012;

Nicholson et al., 2012; Honda and Littman, 2016; Sharon et al., 2016). As a strategic partner of the human holobiont, the GM expands our metabolic potential by adding degradative functions, thus enabling the metabolism of a wide range of complex polysaccharides, otherwise indigestible, including exogenous carbohydrates introduced with diet (Schnorr et al., 2014). Indeed, despite the wide distribution of complex polysaccharides in edible plants (Lattimer and Haub, 2010), our genome is astonishingly poor in genes coding for carbohydrate-active enzymes (CAZymes) (Gill et al., 2006), i.e., enzymes specifically devoted to polysaccharide degradation. Conversely, the gut microbiome (i.e., the cumulative genome of the GM) encodes for a much wider and diversified arsenal of CAZymes (El Kaoutari et al., 2013), allowing for the catabolism of the vast array of complex polysaccharides that reach the colon undigested. These molecules are converted to short-chain fatty acids (SCFAs), microbial metabolites with multiple roles in human physiology (Koh et al., 2016). This inter-kingdom cross-feeding is the basis of the mutualistic relationship we share with our intestinal microbial counterpart, which thus emerges as a key player in the co-metabolism of complex carbohydrates in our gut. In this scenario, the dietary polysaccharides that are metabolically available to gut microbes have been specifically defined as microbiota-accessible carbohydrates (MACs) (Sonnenburg and Sonnenburg, 2014) and their abundance in the host diet, and consequently their availability within the gut, has proved to be crucial for microbiota–host homeostasis.

The MAC availability in the host diet is among the aspects that have deeply changed in recent human evolutionary history, with the transition from Paleolithic hunting-gathering to Neolithic rural populations and contemporary Westernized societies (Quercia et al., 2014; Schnorr et al., 2014; Sonnenburg and Sonnenburg, 2014; Obregon-Tito et al., 2015). In particular, along with Westernization, we have witnessed a progressive reduction of the MAC content and diversity in the diet, with the transition to foods high in refined simple sugars. According to Sonnenburg et al. (2016), this reduction of dietary MACs forced the progressive impoverishment of the CAZyme repertoire in the human microbiome, compromising the overall metabolic plasticity of the human meta-organism. This was confirmed by gut metagenome studies comparing the CAZyme repertoire between traditional populations consuming high-fiber diets and Western urban citizens, which highlighted a relevant reduction of the microbial CAZyme diversity in the latter (King et al., 2012; Bhattacharya et al., 2015; Rampelli et al., 2015). Such wide GM comparative surveys provided important information on the GM–host co-evolutionary dynamics but they did not allow for a precise discrimination of the impact of individual covariates (e.g., diet, health, medication) on the GM functional repertoire. Aiming to fill this gap in knowledge, here we compared the imputed CAZyme repertoire from two cohorts of Western urban adults from Italy, overcoming the intrinsic variation of the CAZyme profiles according to geography and ethnicity (Soverini et al., 2016). In particular, we analyzed the GM structure and metadata from 16 healthy subjects following a Mediterranean diet and 40 obese, type 2 diabetic (T2D) patients consuming a high-fat low-MACs diet

from two previously published studies (Schnorr et al., 2014; Candela et al., 2016). Our findings highlighted the existence of a potentially limited number of well-balanced host–microbe symbiotic configurations, with a possible connection to diet and health status.

MATERIALS AND METHODS

Determination of the Pan-microbiome from Italian Healthy Subjects

The publicly available 16S rRNA sequencing data of the fecal samples of 16 Italian healthy subjects from Schnorr et al. (2014) were downloaded from the MG-RAST website¹ and taxonomically characterized to the species level using the QIIME pipeline (Caporaso et al., 2010), with blastn as an assignment method and the HMP gastrointestinal 16S rRNA dataset as reference sequences². The detected species were considered part of the so-called Italian “pan-microbiome,” i.e., the virtual entity gathering the vast majority of bacterial species present in the GM of the Italian population. The assembled reference genomes of these bacterial species were downloaded from the NCBI genome section³. Then, to characterize the CAZyme repertoire of these microorganisms, the CAZyme identification pipeline developed by Soverini et al. (2016) was applied. Briefly, ORFs were extracted from the assembled genomes using FragGeneScan 1.16 (Rho et al., 2010). From the translated ORFs, the CAZyme-coding sequences were detected using the hmmscan tool of the HMMER software package (Robert et al., 2011) and the dbCAN CAZyme database (Yin et al., 2012). The outputs were further processed by the script hmmscan-parser.sh⁴, selecting only the ORFs that showed a minimum identity of 30% to the query sequences and an alignment length of at least 100 residues.

Identification of CAZyme Co-abundance Groups within the Italian Pan-microbiome

The CAZyme profiles were used to generate CAZy co-abundance groups (CCGs), which were conceived as groups of bacterial species sharing a similar CAZyme profile. In brief, the CCGs were generated by applying hierarchical Ward-linkage clustering based on Spearman correlation coefficients to the abundances of glycosyl-hydrolase (GH) and auxiliary activity (AA) families detected in the bacterial genomes. Permutational multivariate analysis of variance (function “adonis” of the vegan package in R) was used to determine whether CCGs were significantly different from each other. CAZymes were also manually classified for their ability to degrade specific substrates by consulting the publicly available CAZy database⁵. Specifically, we evaluated the ability to degrade different types of MACs: resistant starch (RS),

¹<http://metagenomics.anl.gov/linkin.cgi?project=mgp8810>

²<http://hmpdacc.org/HM16STR/healthy/>

³<https://www.ncbi.nlm.nih.gov/genome/>

⁴https://github.com/carden24/Bioinformatics_scripts/blob/master/hmmscan-parser.sh

⁵<http://www.cazy.org>

non-digestible carbohydrates (NDC), non-starch polysaccharides (NSP), and mucins/glycoproteins (M/G). When more than one activity was found, we selected the most relevant one, i.e., the one with the highest abundance of genes involved in the degradation of a given substrate.

Assessment of Redundant Patterns of CAZymes in Italian Healthy Subjects and Type 2 Diabetes Patients

To explore CAZyme profiles in the Italian population in health and disease, we integrated the dataset used to determine the pan-microbiome with the 16S rRNA sequences of the GM from 40 patients affected by type 2 diabetes (Candela et al., 2016). The sequences were downloaded from MG-RAST⁶ and analyzed using QIIME (Caporaso et al., 2010) and the HMP database, as described above for healthy subjects. The CAZyme profile of each GM was obtained by quantifying the relative abundance of each CCG, as a sum of the relative contribution of component bacterial species. We then grouped the subjects using hierarchical Ward-linkage clustering based on Spearman correlation coefficients. Separation between clusters was tested using the permutational multivariate analysis of variance (function “adonis” of vegan). All statistical analyses were computed in R version 3.1.3 using R studio version 1.0.36 with packages vegan and made4.

RESULTS

The Italian Pan-microbiome from Healthy Adults

The analysis of the publicly available 16S rRNA sequences of stool samples from 16 Italian healthy adults (aged 20–40 years, mean 32 years) consuming a standard Mediterranean diet (Schnorr et al., 2014) led to the identification of a total of 98 bacterial species, present at least once in the samples (Supplementary Table S1). *Faecalibacterium prausnitzii* was the most represented species, with an average relative abundance of $11 \pm 0.09\%$ (standard deviation of the mean). *Eubacterium rectale* ($7 \pm 0.06\%$), *Ruminococcus bromii* ($6 \pm 0.05\%$), and *Bifidobacterium adolescentis* ($6 \pm 0.06\%$) occurred as co-dominant species. *Subdoligranulum variabile* ($3 \pm 0.04\%$), *Ruminococcus champanellensis* ($3 \pm 0.02\%$), *Clostridium asparagiforme* ($2 \pm 0.01\%$), *Bacteroides vulgatus* ($2 \pm 0.03\%$), *Coprococcus eutactus* ($2 \pm 0.03\%$), and *Roseburia intestinalis* ($1.8 \pm 0.01\%$) were ancillary species present to a lower extent. In terms of prevalence, *F. prausnitzii*, *E. rectale*, *Butyrivococcus pullicaecorum*, *R. intestinalis*, *Ruminococcus* spp., *Flavonifractor plautii*, *Blautia obeum*, *Dorea formicigenerans* and *Anaerostipes hadrus* were present in all samples of the dataset, whereas *Methanobrevibacter smithii*, *Bifidobacterium catenulatum*, *Lactobacillus ruminis*, *Weissella paramesenteroides*, and *Sutterella parvibrubra* were found in less than three samples, being among the least prevalent microorganisms in the analyzed Italian microbiomes.

⁶<http://metagenomics.anl.gov/mgmain.html?mgpage=project&project=mgp17675>

CAZyme Repertoire in the Bacterial Species of the Italian Pan-microbiome

Reference genomes from the bacterial species included in the Italian pan-microbiome were retrieved from the NCBI database, and the respective CAZyme-coding sequences were identified. The number of these sequences varied widely between the pan-microbiome components. In particular, *Odoribacter laneus*, *Lactobacillus salivarius*, and *F. plautii* showed the lowest amounts of CAZyme-coding genes, while *Eubacterium bifforme*, *Dialister succinatiphilus*, and *D. formicigenerans* the highest (see Supplementary Figure S1 for retrieving the number of CAZyme-coding sequences for each bacterial species).

To identify common patterns of MAC-degrading enzymes among the different species of the Italian pan-microbiome, we determined co-abundance associations between the bacterial CAZyme profiles and then clustered them based on similarity. Four robust CAZy Co-abundance Groups (CCGs) were identified in the Italian pan-microbiome, each one including GM species that share a similar CAZyme profile, describing the pattern of CAZyme variation within the Italian GM species (p -value < 0.001 , permutation test with pseudo F ratios). CCGs were named according to the most abundant species in each group as follows: *S. variabile* (CCG1), *E. rectale* (CCG2), *R. bromii* (CCG3), and *F. prausnitzii* (CCG4) (Figure 1). Interestingly, *F. prausnitzii* and *R. bromii* CCGs included species with the highest amount of CAZyme-coding sequences, such as *D. succinatiphilus*, *D. formicigenerans*, and *E. bifforme*. On the other hand, CCG1 (*S. variabile*) and CCG2 (*E. rectale*) included species with a lower amount of CAZyme genes, such as *O. laneus* and *F. plautii*. It is important to note that the bacterial genome size is not predictive of the number of encoded CAZymes (El Kaoutari et al., 2013). For example, *Bacteroides oleiciplenus*, *Bacterioides cellulosilyticus* and *Bacteroides ovatus*, the three species with the largest genome size in the Italian pan-microbiome (Supplementary Figure S1), were found to be included within the *S. variabile* group and show the lowest richness in the CAZyme repertoire.

Finally, we specifically evaluated the distribution of CAZyme families involved in the degradation of different types of MACs across the four CCGs (Figure 2). According to our findings, *E. rectale* and *F. prausnitzii* groups showed the highest number of sequences encoding for CAZymes involved in the degradation of non-digestible carbohydrates and non-starch polysaccharides, while the *R. bromii* group was the least rich in these enzymatic functions (p -value < 0.05 , Wilcoxon rank sum test). Interestingly, the ability to degrade RS was evenly spread among the different CCGs.

CAZyme Distribution in the Gut Microbiome of Italian Individuals: Associations with the Diet and Health Status

We next explored the variation of CCG profiles in Italian subjects, also in association with the dietary pattern and health conditions. To this aim, the dataset of analyzed individuals was integrated

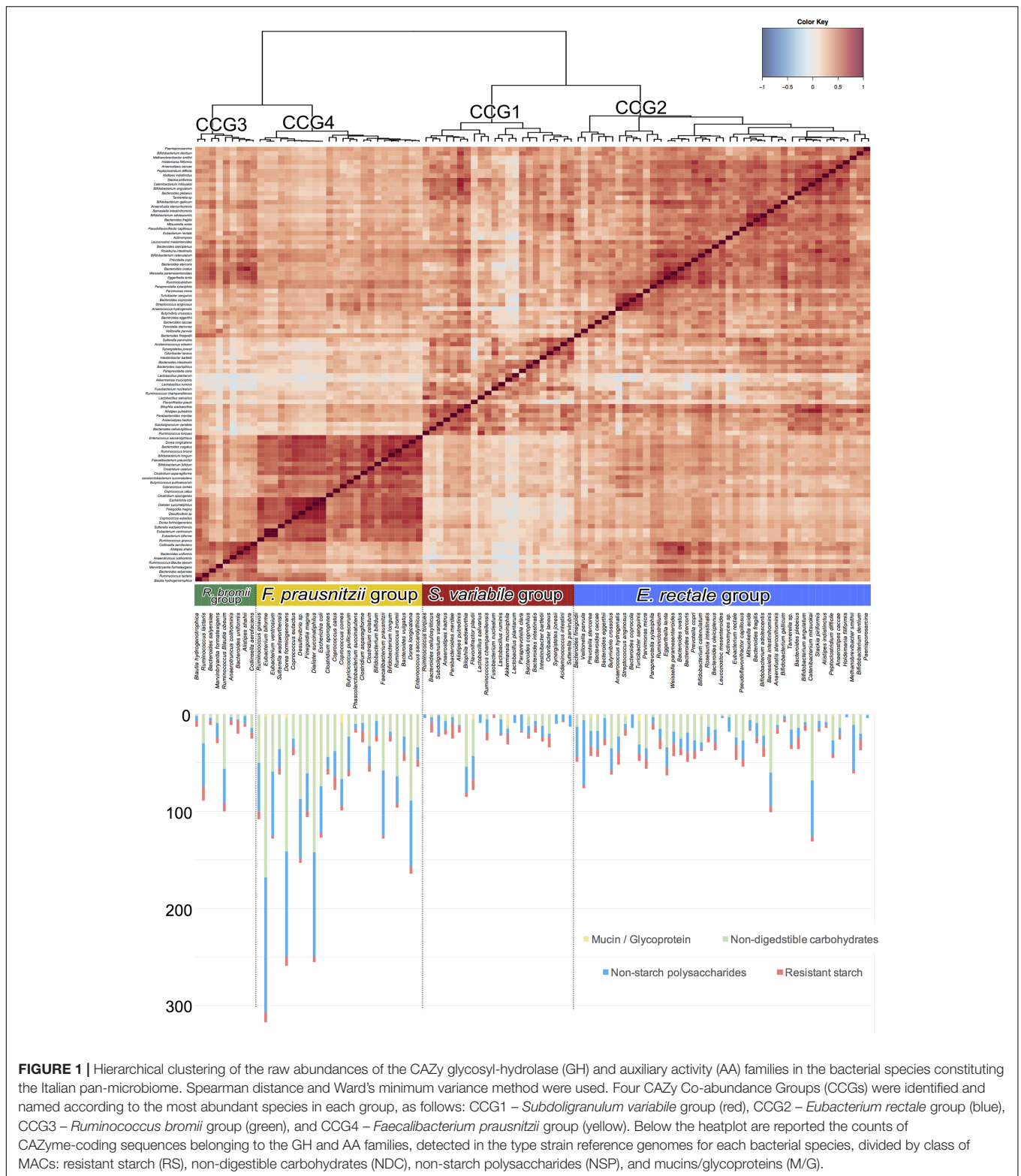
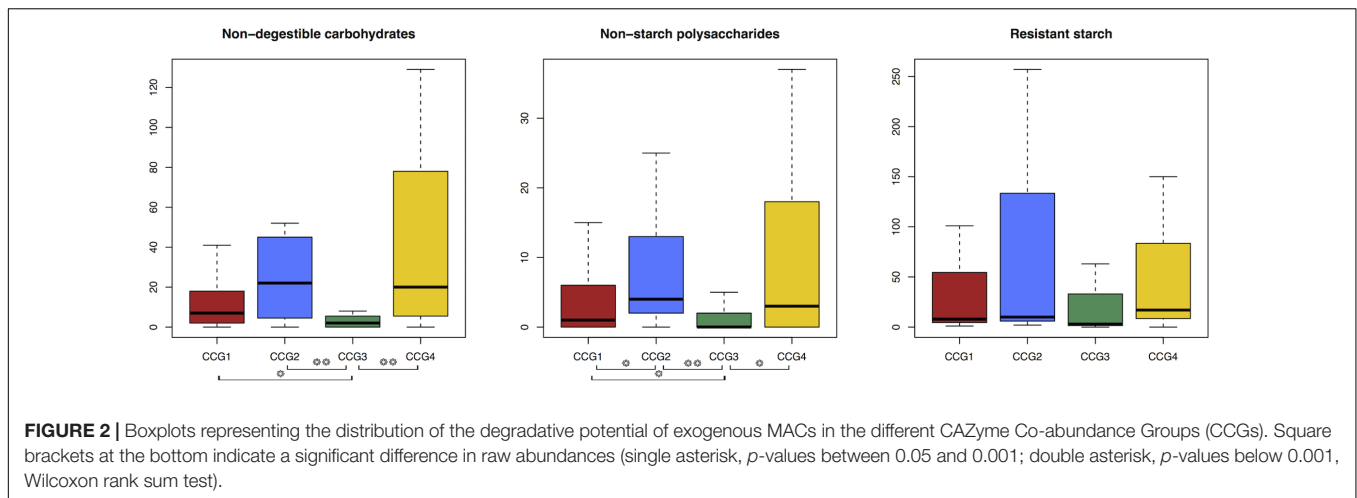


FIGURE 1 | Hierarchical clustering of the raw abundances of the CAZy glycosyl-hydrolase (GH) and auxiliary activity (AA) families in the bacterial species constituting the Italian pan-microbiome. Spearman distance and Ward’s minimum variance method were used. Four CAZy Co-abundance Groups (CCGs) were identified and named according to the most abundant species in each group, as follows: CCG1 – *Subdoligranulum variable* group (red), CCG2 – *Eubacterium rectale* group (blue), CCG3 – *Ruminococcus bromii* group (green), and CCG4 – *Faecalibacterium prausnitzii* group (yellow). Below the heatmap are reported the counts of CAZyme-coding sequences belonging to the GH and AA families, detected in the type strain reference genomes for each bacterial species, divided by class of MACs: resistant starch (RS), non-digestible carbohydrates (NDC), non-starch polysaccharides (NSP), and mucins/glycoproteins (M/G).

with publicly available 16S rRNA gene sequences from 40 Italian obese patients suffering from type 2 diabetes and consuming a high-fat low-MACs diet (Candela et al., 2016). The addition of these samples did not result in an increase in the size of the

Italian pan-microbiome, which still consisted of the previously identified 98 species. To compare the CCG distribution among individual gut microbiomes, for each CCG, the sum of the abundances of constituent microorganisms was calculated, and



the personal profile of CCG abundance in the GM was computed for each subject. According to a clustering analysis of the individual CCG abundance profiles, subjects could be clustered into four groups (p -value < 0.001, permutation test with pseudo F ratio), each including individuals that shared a comparable profile of the four CCGs, and therefore potentially a similar carbohydrate-degrading functional pattern. We defined these redundant patterns as CAZyTypes (CTs), i.e., clusters of different GM configurations with a similar carbohydrate-degrading profile (Figure 3). Interestingly, CT2 and CT4 were more represented in healthy individuals consuming a Mediterranean diet, while CT1 and CT3 occurred with higher frequency in obese T2D patients consuming a high-fat low-MACs diet. By Principal Coordinates Analysis (PCoA), we represented the CCG variation patterns in individual microbiomes in a two-dimensional space (Supplementary Figure S2), highlighting the separation of subjects based on both the CCG abundance and health status (p -value < 0.001 permutation test with pseudo F ratio).

Looking at the CCG distribution in the Italian population, we found that the most prevalent CCG was *F. prausnitzii* group, with the co-presence of *E. rectale* group as an ancillary CCG in the health-prevalent CT2 and CT4 (Figure 3). On the other hand, *S. variable* group and *R. bromii* group were sporadically represented in the subjects of all CTs, but generally present in lower abundance in healthy subjects. Indeed, only three healthy subjects out of 16 showed a relative abundance of the *S. variable* and *R. bromii* groups higher than the average contribution in the overall dataset. Conversely, a higher abundance of these two CCGs was found in T2D patients (18 out of 40).

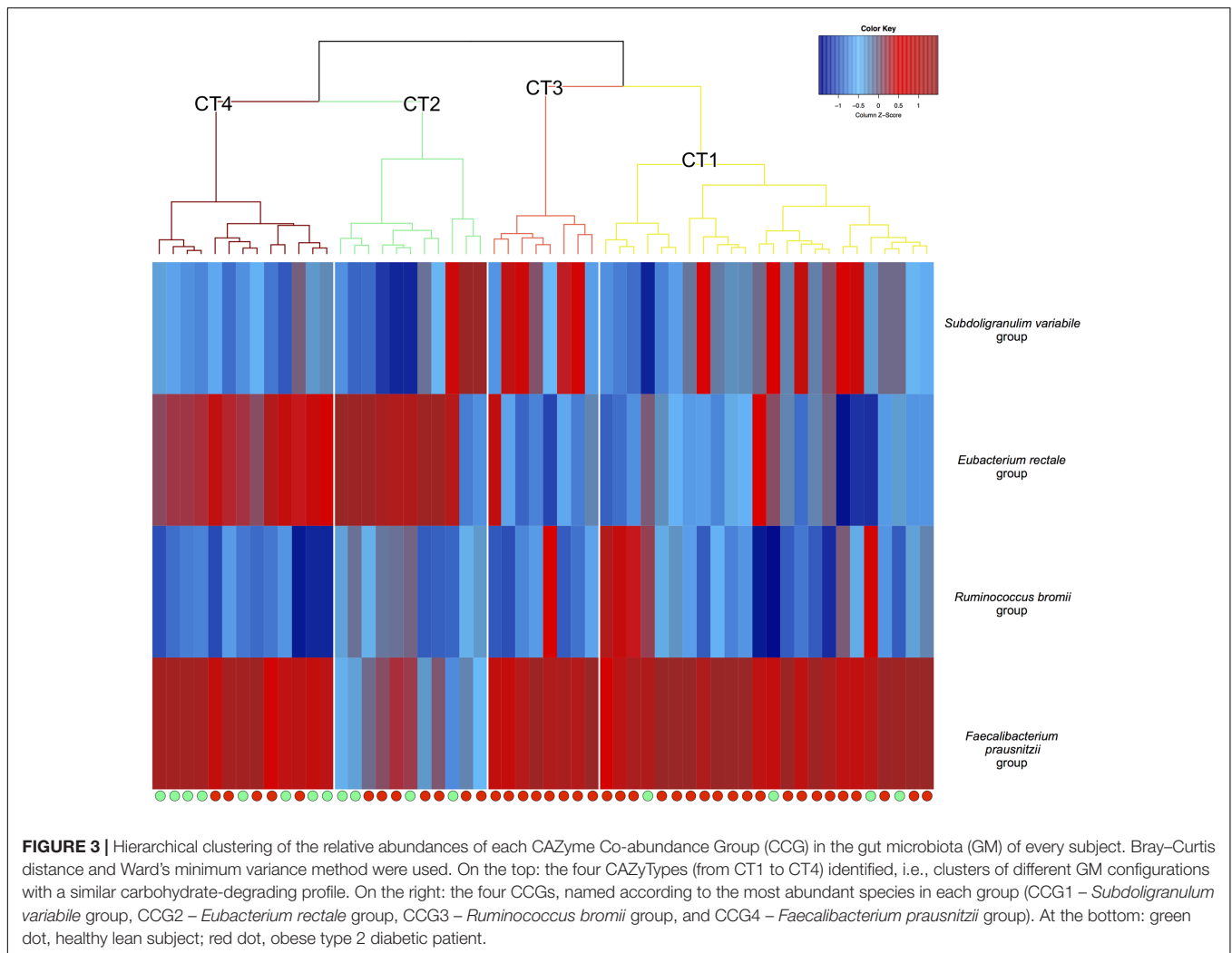
DISCUSSION

Our work aimed at characterizing the frame of the CAZyme variation in the GM meta-community from 56 Italian subjects, of whom 16 were healthy lean adults consuming a standard Mediterranean diet (Schnorr et al., 2014) and 40 were obese T2D patients consuming a high-fat low-MACs diet (Candela et al., 2016). The pan-microbiome of the studied population

included a total of 98 bacterial species and, in healthy subjects, it was dominated by *F. prausnitzii*, *E. rectale*, *R. bromii* and *B. adolescentis*, which also emerged as the most prevalent species in the GM from Italian healthy adults. When compared with previously characterized GM pan-genomes, such as that from the Chinese population (Zhang et al., 2015), the Italian one showed some peculiarities, i.e., the presence of *E. rectale* and *Bifidobacterium* and the absence of *Phascolarctobacterium* within the core community. Although both studies are based on relatively small cohorts, and a more extensive screening at the population level is needed, these data seem to suggest a certain level of country specificity in the gut microbiome structure, which may contribute to the immunological and metabolic peculiarities of the populations.

According to our findings, the bacterial species belonging to the Italian pan-microbiome showed two different types of MAC-degrading profiles, essentially characterized by a high or low content of glycosyl-hydrolase-coding sequences, respectively. As expected, the CAZyme distribution in the various species of the Italian GM was heterogeneous, and the absolute number of CAZymes was independent from the genome size. However, it should be mentioned that, for each identified species, the analysis of the CAZyme content was performed by using the type strain reference genome deposited in the NCBI database and therefore our classification was blind with respect to the possible strain-level functional variability in the CAZyme profile.

The GM species of the Italian pan-microbiome were successfully clustered into four CCGs according to the similarity of the CAZyme pattern: *S. variable* (CCG1), *E. rectale* (CCG2), *R. bromii* (CCG3), and *F. prausnitzii* (CCG4). Interestingly, each of the identified CCGs was characterized by a peculiar structure in terms of CAZyme content. In particular, *F. prausnitzii* and *R. bromii* groups were the most enriched CCGs in terms of represented CAZyme functions, whereas *F. prausnitzii* and *E. rectale* groups were the most equipped CCGs in terms of CAZymes specifically involved in the breakdown of non-digestible carbohydrates and non-starch polysaccharides (i.e., xylans, pectins, and mannans). These observations suggest that the Italian pan-microbiome is diversified in at least



four patterns of carbohydrate degradation, raising several open questions related to: (i) the major determinants of the co-evolutionary processes underlying this differentiation; (ii) the relative contribution of host genetics, lifestyle and diet as drivers of this functional convergence; (iii) the ultimate connections between the observed CCGs and the host metabolic phenotype.

When exploring the quantitative distribution of CCGs in the individual microbiota from all subjects analyzed, we observed four robust clusters of subjects sharing a similar CCG profile, termed from CT1 to CT4. In particular, the CT1 and CT3 clusters included CCG4 (*F. prausnitzii* group) as the most prevalent CCG, being present in all individuals, and CCG3 (*R. bromii* group) and CCG1 (*S. variabile* group) as less prevalent, ancillary, and generally mutually exclusive groups. Conversely, CT4 was dominated by both CCG4 (*F. prausnitzii* group) and CCG2 (*E. rectale* group), which equally shared the ecosystem. Finally, CT2 showed CCG2 (*E. rectale* group) as the most prevalent group and CCG4 (*F. prausnitzii* group) as ancillary and less prevalent group, except for three subjects that were dominated by CCG1 (*S. variabile* group). These observations are indicative

of a different ecological behavior for the diverse CCGs. Indeed, while CCG4 (*F. prausnitzii* group) appears to co-exist with all other CCGs, CCG2 (*E. rectale* group) and the CCGs 1 (*S. variabile* group)/3 (*R. bromii* group) are mutually exclusive. Confirming this, none of the CTs showed the simultaneous presence of CCG2 and CCG1 and/or CCG3. Taken together, these data suggest that the GM-host co-evolution process has resulted in the establishment of four well-defined functional steady states, i.e., the four CTs, each determined by the CCG propensity to share the same gut environment, and each conferring to the host a specific pattern of CAZymes.

In order to explore possible associations of these CTs with the host diet and health, we explored their variation in Italian healthy adults consuming a Mediterranean diet and obese T2D patients consuming a high-fat low-MACs diet. Interestingly, according to our data, most healthy subjects belonged to the CTs 2 and 4, which were characterized by the simultaneous presence of CCG4 (*F. prausnitzii* group) and CCG2 (*E. rectale* group). Conversely, the great majority of obese T2D patients belonged to CT1 and CT3, where CCG2 (*E. rectale* group)

was substituted by CCG3 (*R. bromii* group) and/or CCG1 (*S. variabile* group). Although caution is needed in interpreting results, the analysis presented here suggests that a high-fat low-MACs diet, in the context of metabolic deregulation, such as obesity and T2D, could force changes in the GM CTs, supporting the presence of CCG1 (*S. variabile* group) and/or CCG3 (*R. bromii* group) to the detriment of CCG2 (*E. rectale* group). Interestingly, compared to CCG1 and CCG3, the CCG2 showed higher levels of enzymes involved in the degradation of non-digestible carbohydrates and non-starch polysaccharides, which are indeed abundant MACs in the Mediterranean dietary regimen. Though preliminary, our data highlight a possible adaptive or maladaptive nature for each of the four CT steady states that describe the Italian pan-microbiome. Indeed, the steady states CTs 2 and 4, that were generally associated with healthy hosts, seem to be the result of an adaptive microbiome-host co-evolution process, in which the interplay between diet, gut microorganisms and the host can contribute to overall metabolic health. On the contrary, the CTs 1 and 3 that were associated with T2D and a high-fat low-MACs diet, may result from a maladaptive microbiome-host process, in which this type of diet has led to the selection of CT steady states able to contribute to metabolic and/or immunological deregulation. Supporting these hypotheses, patients suffering from Behçet's syndrome, a systemic inflammatory condition, showed a specific GM functional dysbiosis accompanied by a decreased production of SCFAs (Consolandi et al., 2015). Although the CAZyme profile has not yet been explored in patients with this disease, we can speculate that the observed reduction in SCFA biosynthesis could be the result of a maladaptive transition to CT1 or CT3, which would

diminish the GM potential to provide the host with essential metabolites, such as butyrate, crucial to support immunological health.

CONCLUSION

Our findings highlighted the existence of specific and well-defined GM functional layouts (CAZyTypes, CTs) for what concerns the ecosystem capacity to metabolize MACs, and support the hypothesis that the human GM has the ability to reconfigure its own CAZyme functional layout in response to dietary changes, with possible implications for the host health and metabolic regulation.

AUTHOR CONTRIBUTIONS

SR and MS: conceived, designed and performed the analysis; SR, MS, MC, and ST: wrote the manuscript; EB, SQ, and PB: revised and edited the draft. All authors discussed the results, commented on the manuscript and approved the final version.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2017.02079/full#supplementary-material>

TABLE S1 | Prevalence data and average abundance of the identified 98 bacterial species in the dataset.

REFERENCES

- Bhattacharya, T., Ghosh, T. S., and Mande, S. S. (2015). Global profiling of carbohydrate active enzymes in human gut microbiome. *PLoS ONE* 10:e0142038. doi: 10.1371/journal.pone.0142038
- Candela, M., Biagi, E., Soverini, M., Consolandi, C., Quercia, S., Severgnini, M., et al. (2016). Modulation of gut microbiota dysbioses in type 2 diabetic patients by macrobiotic Ma-Pi 2 diet. *Br. J. Nutr.* 116, 80–93. doi: 10.1017/S0007114516001045
- Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., et al. (2010). QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* 7, 335–336. doi: 10.1038/nmeth.f.303
- Consolandi, C., Turrioni, S., Emmi, G., Severgnini, M., Fiori, J., Peano, C., et al. (2015). Behçet's syndrome patients exhibit specific microbiome signature. *Autoimmun. Rev.* 14, 269–276. doi: 10.1016/j.autrev.2014.11.009
- El Kaoutari, A., Armougom, F., Gordon, J. I., Raoult, D., and Henrissat, B. (2013). The abundance and variety of carbohydrate-active enzymes in the human gut microbiota. *Nat. Rev. Microbiol.* 11, 497–504. doi: 10.1038/nrmicro.3050
- Gill, S. R., Pop, M., Deboy, R. T., Eckburg, P. B., Turnbaugh, P. J., Samuel, B. S., et al. (2006). Metagenomic analysis of the human distal gut microbiome. *Science* 312, 1355–1359. doi: 10.1126/science.1124234
- Honda, K., and Littman, D. R. (2016). The microbiota in adaptive immune homeostasis and disease. *Nature* 535, 75–84. doi: 10.1038/nature18848
- Hooper, L. V., Littman, D. R., and Macpherson, A. J. (2012). Interactions between the microbiota and the immune system. *Science* 336, 1268–1273. doi: 10.1126/science.1223490
- King, D. E., Mainous, A. G., and Lambourne, C. A. (2012). Trends in dietary fiber intake in the United States, 1999–2008. *J. Acad. Nutr. Diet.* 112, 642–648. doi: 10.1016/j.jand.2012.01.019
- Koh, A., De Vadder, F., Kovatcheva-Datchary, P., and Bäckhed, F. (2016). From dietary fiber to host physiology: short-chain fatty acids as key bacterial metabolites. *Cell* 165, 1332–1345. doi: 10.1016/j.cell.2016.05.041
- Lattimer, M. J., and Haub, M. D. (2010). Effects of dietary fiber and its components on metabolic health nutrients. *Nutrients* 2, 1266–1289. doi: 10.3390/nu2121266
- Neish, A. S. (2009). Microbes in gastrointestinal health and disease. *Gastroenterology* 136, 65–80. doi: 10.1053/j.gastro.2008.10.080
- Nicholson, J. K., Holmes, E., Kinross, J., Burcelin, R., Gibson, G., Jia, W., et al. (2012). Host-gut microbiota metabolic interactions. *Science* 336, 1262–1267. doi: 10.1126/science.1223813
- Obregon-Tito, A. J., Tito, R. Y., Metcalf, J., Sankaranarayanan, K., Clemente, J. C., Ursell, L. K., et al. (2015). Subsistence strategies in traditional societies distinguish gut microbiomes. *Nat. Commun.* 25:6505. doi: 10.1038/ncomms7505
- Quercia, S., Candela, M., Giuliani, C., Turrioni, S., Luiselli, D., Rampelli, S., et al. (2014). From lifetime to evolution: timescales of human gut microbiota adaptation. *Front. Microbiol.* 5:587. doi: 10.3389/fmicb.2014.00587
- Rampelli, S., Schnorr, S. L., Consolandi, C., Turrioni, S., Severgnini, M., Peano, C., et al. (2015). Metagenome sequencing of the Hadza hunter-gatherer gut microbiota. *Curr. Biol.* 25, 1682–1693. doi: 10.1016/j.cub.2015.04.055
- Rho, M., Tang, H., and Ye, Y. (2010). FragGeneScan: predicting genes in short and error-prone reads. *Nucleic Acids Res.* 38:e191. doi: 10.1093/nar/gkq747

- Robert, D., Finn, J. C., and Sean, R. E. (2011). HMMER web server: interactive sequence similarity searching. *Nucleic Acids Res.* 39, W29–W37. doi: 10.1093/nar/gkr367
- Schnorr, S. L., Candela, M., Rampelli, S., Centanni, M., Consolandi, C., Basaglia, G., et al. (2014). Gut microbiome of the Hadza hunter-gatherers. *Nat. Commun.* 15:3654. doi: 10.1038/ncomms4654
- Sharon, G., Sampson, T. R., Geschwind, D. H., and Mazmanian, S. K. (2016). The central nervous system and the gut microbiome. *Cell* 167, 915–932. doi: 10.1016/j.cell.2016.10.027
- Sonnenburg, E. D., Smits, S. A., Tikhonov, M., Higginbottom, S. K., Wingreen, N. S., and Sonnenburg, J. L. (2016). Diet-induced extinctions in the gut microbiota compound over generations. *Nature* 534, 212–215. doi: 10.1038/nature16504
- Sonnenburg, E. D., and Sonnenburg, J. L. (2014). Starving our microbial self: the deleterious consequences of a diet deficient in microbiota-accessible carbohydrates. *Cell Metab.* 20, 779–786. doi: 10.1016/j.cmet.2014.07.003
- Soverini, M., Rampelli, S., Turrone, S., Schnorr, S. L., Quercia, S., Castagnetti, A., et al. (2016). Variations in the post-weaning human gut metagenome profile as result of *Bifidobacterium* acquisition in the western microbiome. *Front. Microbiol.* 12:1058. doi: 10.3389/fmicb.2016.01058
- Yin, Y., Mao, X., Yang, J. C., Chen, X., Mao, F., and Xu, Y. (2012). dbCAN: a web resource for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res.* 40, 445–451. doi: 10.1093/nar/gks479
- Zhang, J., Guo, Z., Xue, Z., Sun, Z., Zhang, M., Wang, L., et al. (2015). A phylo-functional core of gut microbiota in healthy young Chinese cohorts across lifestyles, geography, and ethnicities. *ISME J.* 9, 1979–1990. doi: 10.1038/ismej.2015.11

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Soverini, Turrone, Biagi, Quercia, Brigidi, Candela and Rampelli. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.