UNIVERSITÀ DI PARMA

University of Parma

Department of Chemistry, Life Sciences and Environmental Sustainability

Ph.D. in Biotechnology and Life Science

XXX cycle 2015-2017

Disentangling the role of intestinal bacteria in human and animal health

Ph.D. thesis

Ph.D. Coordinator: Prof. Simone Ottonello

Tutor: Prof. Marco Ventura

Ph.D. student

Leonardo Mancabelli

Parma 2017

Table of Contents

Summary

In the last decade, the advent of the Next-Generation high throughput Sequencing (NGS) technologies allowed to shed light on the composition and functionality of the microbial populations inhabiting humans and other animals, i.e. microbiota. In humans, it has been estimated that 10^{14} microorganisms reside in several compartments of the body such as the surface of skin and in the gastrointestinal, genitourinary and respiratory tracts. The gastrointestinal tract, which has the large numbers of microorganisms in humans, included specific compartments such as mouth, esophagus, stomach, small intestine, large intestine (colon), rectum, and anus. The range of published studies focusing on this topic revealed that the gut microbiota is largely responsible for the overall health status of the host. In detail, the complex activities exerted by the gut microbiota impact on several functions of its host, including gut physiology, intestinal metabolism, and immune system modulation. While the gut microbiota composition is influenced by factors such as diet, lifestyle and environment, the alteration of the gut microbiota composition, i.e. dysbiosis, has been associated with a large array of human disorders and diseases such as inflammatory bowel diseases (IBD) and irritable bowel syndrome (IBS), colorectal cancer (CRC), metabolic diseases and allergic disease.

The major role of the microbiota in defining the health status of the host led to extensive study of farm animals. The gut microbiota of healthy animals is supposed to exert a significant role in nutrient assimilation, vitamin and amino acid production as well as prevention of pathogens colonization. Moreover, the livestock animal gut microbiota may also act as a source of bacterial pathogens that can spread to human beings or to exploit a role as a reservoir of antibiotic-resistance determinants, which can be transferred to other microorganisms including opportunistic pathogens.

The aim of this Ph.D. thesis is to explore the role of intestinal bacteria in human and animal health. In particular, it aims to investigate compositional and functional differences of poultry animals kept under different housing regimes and nutritional circumstances. Furthermore, focusing the interest on human beings, a comparison between all publicly available human gut shotgun metagenomic datasets

corresponding to urbanized and pre-agricultural societies allowed to validate the assumption that urbanization/industrialization processes have significantly influenced the composition and functionality of the human gut microbiome. Additionally, these metagenomic data allowed genome reconstruction of bacterial taxa that seem to have been lost or gained by individuals living in urban-industrialized countries. Consequently, this PhD thesis also evaluates the correlation between human gut microbiota and diseases in order to identify possible microorganisms known as microbial biomarkers associated with a health/disease status.

Chapter 1

General Introduction

A. General features of the gut microbiota

The human and animal body harbor trillions of microbial cells whose synchronized actions are supposed to exert a key role in the host life. Such microbial cell populations spread their highest density in the intestinal compartment where they collectively form a complex microbial community known as the gut microbiota (Lozupone*, et al.*, 2012). The gut microbiota may be constituted by microorganisms belonging to any of the three domains of life, i.e. Archaea, Bacteria and Eukarya, while they also comprise viruses. These microorganisms are known to establish complex trophic relationships with each other and their human/animal host, ranging from symbiosis to parasitism (Ventura*, et al.*, 2009). In the gut microbiota may reside autochthonous, also known as indigenous, and allochthonous or transient microorganisms (Ventura*, et al.*, 2009). Consequently, only a rather small number of (opportunistic) pathogens are considered inhabitants of the gut microbiota, residing unperturbed within the enteric host microbiota, only to become a health risk to the host when the gut ecosystem is altered and the gut microbiota homeostasis becomes disrupted.

The gastrointestinal (GIT) microbiota composition (Figure 1) may be influenced by several environmental parameters, such as availability of nutrients, water activity, pH, oxygen levels/redox state and temperature, allowing many populations to flourish and employing diverse activities while interacting with their environment, including that of their hosts (Ursell*, et al.*, 2012).

The abundant and diverse members of the normal gut microbiota, i.e. microbiota in eubiosis, have specific function in host nutrient metabolism, xenobiotic and drug metabolism, maintenance of structural integrity of the gut mucosal barrier, immunomodulation, and protection against pathogens (Eggesbo*, et al.*, 2003, Huh*, et al.*, 2012, Sevelsted*, et al.*, 2015). In the eubiotic microbiota the microorganisms spread a climax status represented by the establishment of an homeostasis between all its members (Tamboli*, et al.*, 2004). The alteration of the normal gut microbiota triggered by a wide range of factors, i.e. dysbiosis, can cause shifts in this microbiota balance, thus disrupting the gut microbiota homeostasis. Dysbiosis is usually correlated with harmful effects and may have long-

term consequences leading to disorders or diseases, such as inflammatory bowel diseases (IBD), irritable bowel syndrome (IBS), colorectal cancer (CRC) as well as metabolic and allergic diseases (Spor*, et al.*, 2011, Kim, 2015, Matsuoka & Kanai, 2015, Baothman*, et al.*, 2016).

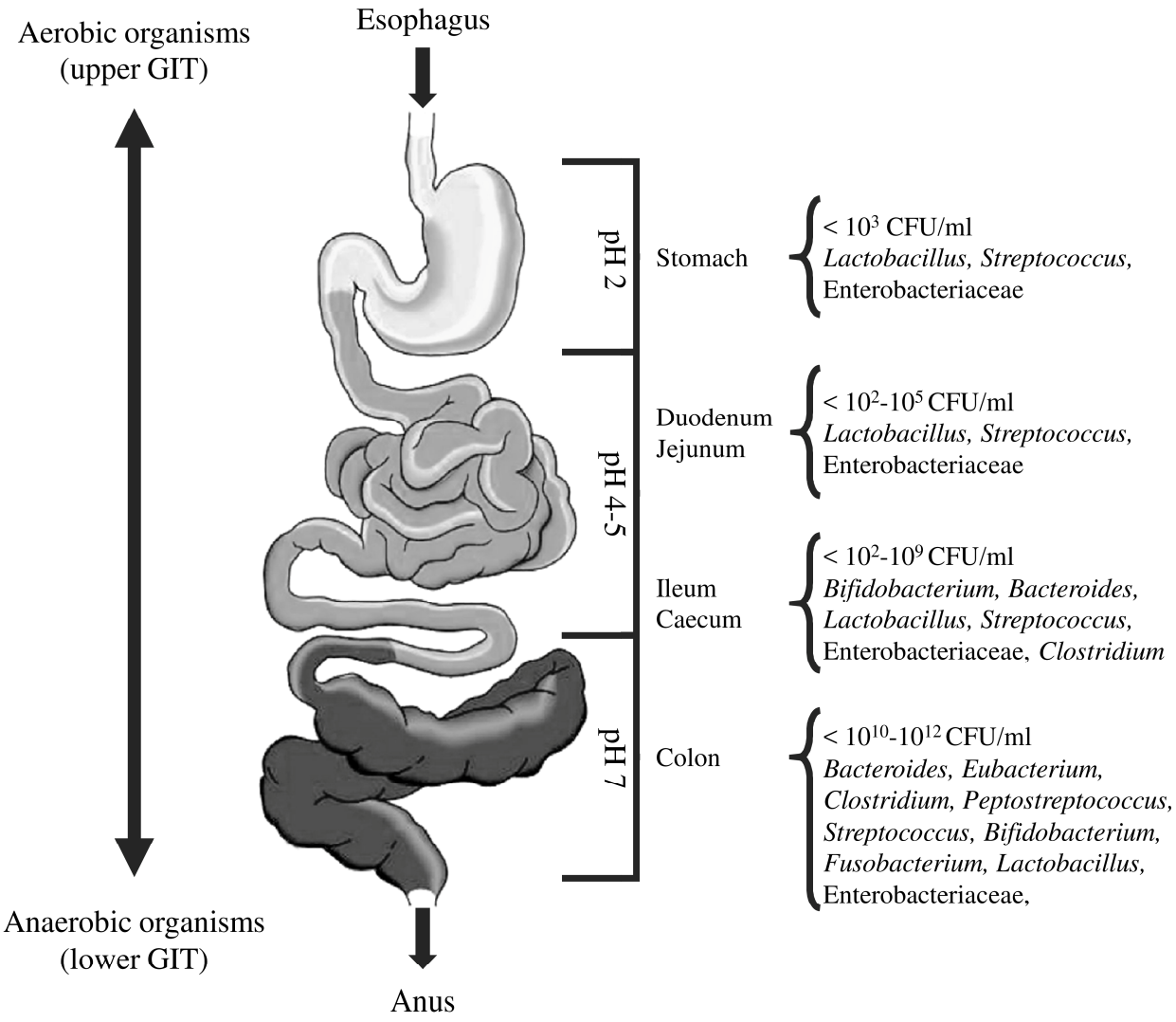


Figure 1: Schematic representation of distribution and abundance of bacteria in human gastrointestinal tract.

Dynamics of the development of the gut microbiota

The microbial communities that populate each individual follow several natural processes, including dispersal, *in situ* diversification, environmental selection and ecological drift, regulating their composition and diversity.

Dispersal is a natural process producing an increase in diversity in local microbial communities, following the assumption that the human body is continually sampling the pool of available microorganisms (Costello*, et al.*, 2012). Another ecological process that influences the microbial communities is local diversification, which provides rapid microbial adaptation via mutation or recombination (Costello*, et al.*, 2012). In this context, horizontal gene transfer (HGT) events may represent one of the chief forces leading microbial variation particularly for those microbes that share the same ecological niche (Smillie*, et al.*, 2011).

Environmental selection is another main ecological process, which molds the human and animal microbiota, and that may be considered as a "habitat filter" consisting of a collection of resources and conditions allowing and/or support growth of certain microorganisms but not others, highlighting the selection of microbial features that let outliving and growth in the host. Remarkably, the total profiles that can originate from environmental selection (niche-based interactions) can differ as a function of the spatial scale across which these processes occur (Kerr*, et al.*, 2002).

Furthermore, the ecological process identified as ecological drift or demographic stochasticity may change microbial taxa abundance. This ecological process is responsible for the disappearance of low-abundance species (e.g., antibiotic-sensitive strains) except they gain a competitive advantage across a different ecological niche or become replenished by dispersal from outside the community (Costello*, et al.*, 2012).

B. Technical approaches for microbiota determination

Although microorganisms are abundant and ubiquitous, few are the knowledge regarding the major roles played by microorganisms in nature, including those residing in the human and animal body (Biteen*, et al.*, 2016). Until the recent development of novel culturomics approaches (Browne*, et al.*, 2016), only a very small portion of the human and animal gut microbiota had been isolated and studied in pure culture (Biteen*, et al.*, 2016). The assumption that a large proportion of the gut microbiota was uncultured (Rajilic-Stojanovic & de Vos, 2014) encouraged the development of culture-independent approaches, such as metagenomics, metatranscriptomics and metaproteomics, in order to discover the identities, activities and functional roles of the so far uncultivated members of the gut microbiota (Figure 2). Currently, the standard methodology for profiling complex microbial communities is based on high-throughput sequencing of a portion of the 16S rRNA gene, i.e. 16S rRNA gene-based microbial profiling analysis, as a conserved phylogenetic marker. The 16S rRNA gene-based microbial profiling approach relies on universal primers for amplification of single or multiple hypervariable regions of the 16S rRNA gene (Hamady & Knight, 2009). The amplicons obtained are sequenced through next-generation sequencing platform and the generated reads are processed using bioinformatic pipelines, such as the popular Qiime software suite (Conde*, et al.*, 2010) or Mothur (Schloss*, et al.*, 2009), so allowing the reconstruction of the microbial composition of the analyzed environmental sample. This procedure also eases identity assignment of unknown members of microbial communities by discrimination based on the sequence of their unique hypervariable regions (Clarridge, 2004). Furthermore, it is possible the use of shotgun metagenomic sequencing strategies to characterize the full genetic content of the gut microbiota, i.e. microbiome (Guinane & Cotter, 2013). However, one of the limits of this metagenomic approach is that the microbiome data does not reveal information on whether or not genes are expressed at any given time. To overcome this limitation other omics approaches have been developed, including the

sequencing of the whole microbial RNA pool of a specified sample, i.e. metatranscriptomics, or analysis of the overall protein content or proteome, i.e. metaproteomics. Notably, the usefulness of these latter two technologies is limited by the fact that many genes or their homologs (and thus their products) are not functionally characterized.

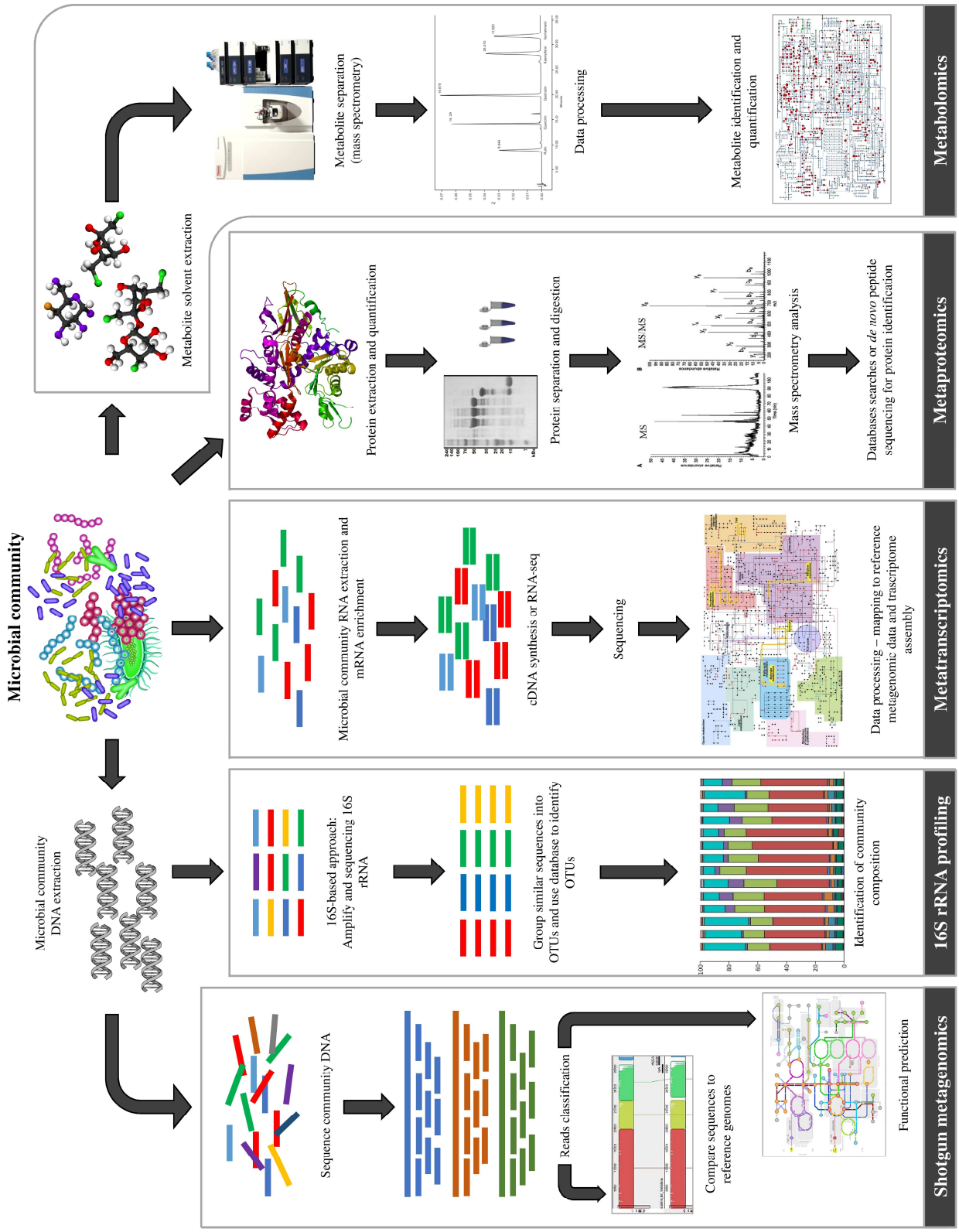


Figure 2: An illustration with the main steps in metagenomic, metatranscriptomic, metaproteomic and metabolomic approaches

Methodologies for the investigation of the microbiota composition.

Several gut microbiota studies were based on 16S rRNA gene-based microbial profiling analyses. The 16S rRNA gene encompasses nine different variable regions, i.e. V1-V9, each flanked by highly conserved DNA sequences that are appropriate for PCR primer binding (Neefs, *et al.*, 1993). Nevertheless, no standard protocols exist to select the most appropriate PCR primer pair that is equally efficient in amplifying part of the 16S rRNA-encoding gene for all taxa and phylotypes present in biological samples, and very often the choice of the appropriated primer pair to use is based on historic practice, anecdotal evidence or/and current literature (Sundquist, *et al.*, 2007, Claesson, *et al.*, 2010, Turroni, *et al.*, 2012, Milani, *et al.*, 2013). Moreover, none of the currently available DNA sequencing technologies offers full-length gene sequencing at adequate depth for cost-effective multiplexing of multiple samples in a single run.

As indicated above, an alternative strategy that allow to profile the gut microbiota is shotgun metagenomic sequencing. This approach bypasses gene-specific amplification and potentially sequences all (fragmented) DNA extracted from the analyzed environmental sample, comprising that from unclassified bacteria and viruses. Shotgun metagenomics reveals more information, including insights into functional aspects of the microbial community, as compared to 16S rRNA gene-based microbial profiling. Interestingly, shotgun metagenomics analysis does not present the potential bias of the amplification reaction required for 16S rRNA gene-based profiling. In detail, shotgun data can be used to investigate the repertoire of genes involved in a wide range of metabolic processes, such as those employed in biosynthesis of compounds, e.g. short chain fatty acids, or in the catabolism of nutrients, e.g. carbon sources. Functional classification of the shotgun metagenomic reads through the use of customized databases may also allow to evaluate antibiotic resistance, degradation of conjugated bile salts, presence of (pro)phages, extracellular structures responsible for adhesion and immunomodulation. Additionally, the complete or partial reconstruction of microorganisms'

genomes through an assembly-based approach may enable the exploration of the microbial dark matter mainly represented by the uncultivated taxa (Rinke*, et al.*, 2013).

Nevertheless, DNA sequencing of complex bacterial communities, such as those residing in the gastrointestinal tract, produces huge amount of data requiring high computational capacity and specific bioinformatics tools for sequence information management, interrogation and administration (Biteen*, et al.*, 2016). Besides, it should be mentioned that underpopulated reference databases and poor functional characterization of many genes considerably limit the usefulness of the metagenomic approaches employed to investigate the gut microbiota.

Approaches to complete a high-definition image of the gut microbiota composition

16S rRNA gene-based microbial profiling analyses provide insights into the composition of complex bacterial communities but its main limit is to be accurate only down to genus level (Chakravorty, *et al.*, 2007). Consequently, in order to overcome this limitation and to acquire a more complete image of the composition of the gut microbiota, i.e. at species or even subspecies level, it is required to target a molecular marker that is much more variable at interspecies level compared to the 16S rRNA gene. A valuable genetic marker for such purpose is represented by the Internally Transcribed Spacer (ITS) sequence, which represents a spacer region between the 16S rRNA and the 23S rRNA genes within the rRNA locus. An ITS-based protocol, i.e. known as ITS-bifidobacterial profiling analysis, was used to attain a detailed image of bifidobacterial communities (Milani, *et al.*, 2014). The ITS-bifidobacterial profiling analysis can discriminate between closely related bifidobacterial taxa, at the subspecies level, and thus to disentangle bifidobacterial community composition in complex ecosystems including the human and animal gut (Milani, *et al.*, 2014, Milani, *et al.*, 2015, Sabbioni, *et al.*, 2016).

A recently useful approach to profile the gut microbiota composition at high resolution down to strain level, encompasses the reconstruction of a genome sequence of an individual microbiota member from shotgun metagenomic data (Lugli, *et al.*, 2017). This metagenomics approach not only reveals taxonomic information about strain identity, but also provides data correlated to the genetic of the organism, thus supplying metabolic and evolutionary insights (Lugli, *et al.*, 2017).

An interesting tool, directed at defining the composition of the gut microbiota at high resolution down to strain level is named MetaPhlAn (Segata, *et al.*, 2012). This software is based on read mapping to a pre-computed database of strain-specific marker genes obtained by means of comparative analysis of all publicly available bacterial genome sequences. The major criticism of this approach is that only

beforehand sequenced species can be profiled, so ignoring the presence of as yet unknown/uncultured members in the population.

Culturomics

During the last decade, the culture-independent approaches have mostly been applied in order to reconstruct the human and animal gut microbiota composition, while microbial cultivation techniques have been partially neglected (Hugon*, et al.*, 2015). This has caused a substantial knowledge gap between bacterial species inhabiting the gastrointestinal tract, but that have not yet been cultivated, and those that have been isolated and cultivated (Hugon*, et al.*, 2015). It has been described that approximately 56 % of gut bacteria detected by metagenomic approaches have cultured representatives (Goodman*, et al.*, 2011, Walker*, et al.*, 2014). With the introduction of so-called culturomic approaches this gap is being closed. Culturomics is based on high-throughput cultivation conditions in order to explore the gut microbiota. Recently, several culturomics studies of human stool samples involved the formulation of complex growth media, which allowed the isolation and cultivation of a large number of novel gut microorganisms (Lagier*, et al.*, 2012, Browne*, et al.*, 2016, Lagier*, et al.*, 2016).

C. Unravel the role of microbiota in health: from human beings to farm animal

Microbiota and diet

Gut microbiota composition is influenced by numerous factors, including host genetics, age (Dicksved, *et al.*, 2008), pregnancy (Koren, *et al.*, 2012) and some environmental factors such as the type of birth (Salminen, *et al.*, 2004), antibiotic intake (Nicholson, *et al.*, 2012) and diet (De Filippo, *et al.*, 2010, David, *et al.*, 2014). In particular, some major research efforts focused on the correlation between intestinal microbiota composition and dietary interventions. Comparison of the human gut microbiota with that of other mammals reveals that the gut microbiota composition of a human being with a modern lifestyle is typical of omnivorous primates (Ley, *et al.*, 2008). It has been revealed that the typical Western diet rich in fat and sugar promotes a shift in the gut microbiota composition towards high numbers of clostridia and a significant decrease in Bacteroidetes (Turnbaugh, *et al.*, 2009). In contrast, a diet containing high levels of fibers promotes an increase of Bacteroidetes and a decrease of Firmicutes (De Filippo, *et al.*, 2010).

Comparison between the fecal microbiota composition of vegetarian *vs.* omnivorous individuals exhibited that the latter are enriched with butyrate-producing bacteria belonging to the *Clostridium* cluster XIVa (Kabeerdoss, *et al.*, 2012, Matijasic, *et al.*, 2014). Conversely, the gut microbiota of vegetarians was shown to be enriched with *Bacteroides* and *Faecalibacterium prausnitzii* species (Dominguez-Bello, *et al.*, 2010). Furthermore, specific studies regarding Western-based diet showed a paucity of fiber-degrading bacteria, such as *Prevotella*, *Succinivibrio*, *Treponema* and bifidobacteria (Yatsunenko, *et al.*, 2012, Ou, *et al.*, 2013, Schnorr, *et al.*, 2014). In contrast, whole grain products have been shown to not only exert a bifidogenic effect with the enrichment of bifidobacteria, but also to rise numbers of *Collinsella*, *Atopobium* and *Clostridium* cluster IV (Carvalho-Wells, *et al.*, 2010).

The comparison of the gut microbiota composition of two human cohorts, receiving either a plant- or meat-based diet, highlighting distinct profiles (David*, et al.*, 2014). In detail, the meat-based diet was shown to increase the abundance of bile-tolerant bacteria such as *Alistipes*, *Bilophila* and *Bacteroides*, with a concomitant decrease of Firmicutes, in particular those members that are involved in the metabolism of plant polysaccharides such as *Roseburia*, *Eubacterium rectale* and *Ruminococcus bromii* (David*, et al.*, 2014).

Focusing the interest on the influence of specific diet-components on the gut microbiota composition, it has observed that humans consuming resistant starch (RS), i.e. dietary fiber, undergo a gut microbiota change with a significant increase in *Ruminococcus bromii* and other starch-degrading bacteria such as *Oscillobacter* and *E. rectale* (Walker*, et al.*, 2005). Dietary fibers also include inulin, which is abundant in many vegetables such as artichoke. Several studies have explored the impact of inulin on the human gut microbiota, and such studies have revealed that a large portion of the gut microbiota remains unaffected by the presence of this glycan, while just a small number of taxa increased in abundance, such as *Bifidobacterium* and *Atopobium* (Costabile*, et al.*, 2010). However, it is not understandable how the level of bifidobacteria increases, since they do not possess the enzymatic arsenal needed for inulin breakdown (Pokusaeva*, et al.*, 2011). However, it is plausible that bifidobacteria cross-feed on inulin degradation products, i.e. fructo-oligosaccharides (FOS), produced by other gut members that possess the required extracellular enzymes for inulin breakdown (Milani*, et al.*, 2015).

The gut microbiota profiling of individual following diets that contain FOS revealed an enrichment of saccharolytic microorganisms, in particular bifidobacteria and *Bacteroides*, with an associated abundance reduction of *Faecalibacterium prausnitzii* and *Roseburia intestinalis* species (Benus*, et al.*, 2010).

Dietary fibers are known to increase growth and activity of butyrate-producing microorganisms such as *Roseburia* sps., *E. rectale* and *F. prausnitzii*. Moreover, bifidobacterial as well as lactobacilli populations are stimulated by the presence of dietary fibers directly or indirectly by cross-feeding.

The influence of fat and protein intake on the gut microbiota composition remains largely unexplained. Preliminary studies have proposed that dietary fat indirectly alters the gut microbiota composition through its impact on bile acid secretion and bile acid composition. In this context, it has been observed that the gut microbiota enriched of *Bacteroides* is positively correlated with the intake of saturated fatty acids (Wu*, et al.*, 2015). Conversely, the gut microbiota enhanced of *Prevotella* genus shows just a weak association with total fat intake (Wu*, et al.*, 2015).

The investigation of the role of the poly-unsaturated fraction of fat indicated its influence on the adherence of intestinal bacteria to mucosal surface (Kankaanpaa*, et al.*, 2001). In particular, animal models fed with a high fat diet displayed a decreased presence of *Bacteroides*-like bacteria, the *E. rectale-Clostridium coccoides* group and members of the genus *Bifidobacterium* (Cani*, et al.*, 2007). Besides, high protein diet has been associated with a lower occurrence of *Bifidobacterium* spp. and with a reduction of the total bacterial counts (Duncan*, et al.*, 2007, Brinkworth*, et al.*, 2009).

Disease and microbiota

As indicated above the gut microbiota interacts with dietary constituents, but it also plays more complex roles such as modulation of the immune system. Moreover, the composition of the gut microbiota may be important in reducing the chance of contracting particular gut infections. Changes in the microbiota over the lifetime of an individual are accompanied by modifications in multiple health parameters, which has prompted intense scientific efforts to understand this complex interaction. Variations in the presence and/or abundance of certain components of the intestinal bacteria have consistently been correlated to particular gut associated diseases, named microbial biomarkers (see below) (Figure 3).

Several studies have been undertaken in order to investigate the relationships between gut microbiota composition and various diseases such as necrotizing enterocolits (Mai*, et al.*, 2011), inflammatory bowel disease (Carroll*, et al.*, 2011, Saulnier*, et al.*, 2011, Matsuoka & Kanai, 2015) and colon cancer (Sobhani*, et al.*, 2011, Gagniere*, et al.*, 2016, Gao*, et al.*, 2017). Most of these diseases have been associated with dysbiosis that might be due to several factors including antibiotic treatment, physical or psychological stresses, radiation, altered peristalsis and dietary shifts (Hawrelak & Myers, 2004).

Figure 3: Schematic representation of the functional roles of key members of the human gut microbiota in health and disease.

Gut microbiota in Inflammatory Bowel Disease (IBD)

Inflammatory bowel disease (IBD) is a group of inflammatory conditions of the colon and small intestine. Crohn's disease (CD) and ulcerative colitis (UC) are the principal types of inflammatory bowel disease. Studies based on germ-free mice demonstrated that experimental colitis does not develop in these animals, whereas the introduction of a single bacterial species can induce mucosal inflammation further implicates gut bacteria in the pathogenesis of IBD. This observation reinforces the hypothesis that intestinal bacteria play an important role in the pathogenesis of IBD. Moreover, reduced bacterial diversity has been demonstrated in mucosal biopsies of patients with active IBD, with loss of commensal species such as *Clostridium leptum*, *Eubacterium* and bifidobacteria (Favier, *et al.*, 1997, Mangin, *et al.*, 2004). Furthermore, the analysis of the fecal microbiota of patients suffering from IBD revealed an under-representation of the Firmicutes phylum and particularly the species *Faecalibacterium prausnitzii* (Sokol, *et al.*, 2008). More specifically, investigation of the CD disease has shown an increase of genera belonging to Proteobacteria phylum in patients with this pathology (Li, *et al.*, 2012). Analysis of samples affected by UC, which are generally characterized by inflammation and ulceration of the lining of the colon, highlighted a positive correlation of this disease to bacteria belonging to the Enterobacteriaceae family (Garrett, *et al.*, 2010, Gerritsen, *et al.*, 2011).

Although microbial sensing has been implicated in the pathogenesis of IBD, recent research has shown that the products of bacterial activity have a regulatory effect on inflammation in IBD (Segain, *et al.*, 2000, Sanderson, 2004). Butyrate is the preferential energy substrate for human colonocytes (Wong, *et al.*, 2006) and it downregulates mucosal inflammatory responses (Segain, *et al.*, 2000, Maslowski, *et al.*, 2009). Patients with CD showed a loss of butyrate-producing species and the actual concentration of butyrate is lower too (van Nuenen, *et al.*, 2004, Manichanh, *et al.*, 2006, Marchesi, *et al.*, 2007, Nemoto, *et al.*, 2012). Moreover, butyrate has been shown to be an effective topical treatment of ulcerative colitis (Scheppach, *et al.*, 1992, Hallert, *et al.*, 2003, Song, *et al.*, 2006). This

reduction suggests that changes in the metabolic activity, as well as composition of intestinal microbiota are important in the pathogenesis of IBD.

Colorectal cancer

Gut microbiota has also been implicated in the development of cancer, especially in colorectal cancer (Davis & Milner, 2009, O'Keefe*, et al.*, 2009, Uronis*, et al.*, 2009). A mechanism by which colonic microbiota may contribute to the onset of colorectal cancer involves the induction of inflammation by commensal bacteria. A recent study involving murine models of colorectal carcinoma mice exposed to carcinogenic agents showed that mice were developing colorectal carcinomas when housed conventionally, but not when housed under germ free conditions (Uronis & Threadgill, 2009). Notably, ingestion of lactic acid bacteria has been noticed to prevent carcinogen-induced lesions as well as tumors in mice (Goldin & Gorbach, 1980, Goldin*, et al.*, 1996, Rowland*, et al.*, 1998).

Although a single causative organism has not been identified, a number of studies have implicated an association for *Fusobacterium* and *Campylobacter* genera with CRC (Brauner*, et al.*, 2010, Castellarin*, et al.*, 2012). In detail, it has been hypothesized *Fusobacterium nucleatum* attaches to the host epithelial E-cadherin and promotes colorectal carcinogenesis via the fusobacterial adhesin FadA (Rubinstein*, et al.*, 2013). Additionally, a recent study demonstrated that *F. nucleatum* promotes colorectal cancer resistance to chemotherapy. In fact, *F. nucleatum* modulates a molecular network of the Toll-like receptor, micro-RNAs, and autophagy to biologically, and mechanistically control colorectal cancer chemoresistance (Yu*, et al.*, 2017). Moreover, Grivennikov et al. investigated the mechanisms responsible for tumor-elicited inflammation in CRC and highlighted in a mouse model of colorectal tumorigenesis an upregulation of interleukin IL-23 and IL-17 that promoted tumor growth, progression and development. In detail, IL-23 is mainly produced by tumor-associated myeloid cells that are likely to be activated by microbial products, which penetrate the tumors but not adjacent tissue. These results indicated that barrier deterioration induced by colorectal-cancer-initiating genetic lesions results in adenoma invasion by microbial products that trigger tumor-elicited inflammation, which in turn drives tumor growth (Grivennikov*, et al.*, 2012).

The microbiota of farm animals

In the last years growing efforts have been directed to analyze the bacterial composition of farm animals, such as chickens, ruminants and swine, representing an efficient commercial/agricultural species. The gut microbiota of healthy livestock animals is presumed to play a key role in nutrient assimilation and prevention of pathogen colonization. Consequently, analysis of the gut microbiota and its management and/or manipulation may help to solve many of the great challenges facing in livestock production, such as minimizing dysbiosis and pathogen colonization, maximizing growth performance and feed efficiency conversion, and maintaining human food safety. Consequently, farm animals with the larger farmed populations in economically developed regions were most studied, such as pigs (Niu*, et al.*, 2015, Heinritz*, et al.*, 2016), cattle (Mao*, et al.*, 2015, Fecteau*, et al.*, 2016, Bessegatto*, et al.*, 2017) and chickens (Choi*, et al.*, 2015, Pourabedin & Zhao, 2015, Wang*, et al.*, 2016).

Pigs and cattle microbiota

Pigs harbor a complex gut microbiota, which establishes strong and complex interactions with the host. Several studies focused their interest toward the composition of the different gastrointestinal section of the pigs, such as ileum, caecum and colon (Figure 4) (Looft*, et al.*, 2014). For example, the ileum lumen samples revealed a lower diversity in terms of bacterial richness and abundance when compared with other intestinal compartments. This comprises almost exclusively Firmicutes and Proteobacteria, whereas the phylum-level profiles of the caecum and mid-colon are highly similar and include mainly Firmicutes, Proteobacteria, Bacteroidetes and Spirochetes (Looft*, et al.*, 2014). Interestingly, mucosa-associated bacterial communities along GIT are different from those present in the lumen. However, statistically significant differences were found solely in the ileum between the mucosal and luminal communities and most lumen-associated bacteria were also found at mucosal level. Moreover, metagenomics analyses of fecal pigs microbiota showed a predominance of the orders Clostridiales, belonging to unclassified Firmicutes, Bacteroidales, Spirochaetales, unclassified Gammaproteobacteria and Lactobacillales. Moreover, functional investigation of the pigs microbiome highlighted significant difference between animals depending on the breed, diet, health status, age and environment (Kim*, et al.*, 2012, Arnal*, et al.*, 2014, Looft*, et al.*, 2014). These evidences suggested that the investigation of pig's gut microbiota is a powerful and versatile tool to predict effects of new feeding/breeding strategies.

Conversely, even if dairy and beef cattle play a central role in dietary and economy of human beings only few studies have assessed their gut microbiota composition using metagenomics approaches.

The currently existing 16S rRNA gene sequencing studies targeting these latter ecosystem revealed a predominance of phyla Firmicutes and Bacteroidetes, followed by the Proteobacteria phylum (Gomez*, et al.*, 2017). In contrast, the phyla Tenericutes and Verrucomicrobia were detected at lower abundance. Furthermore, the genus *Prevotella* was the most abundant taxon and may correlated to

the typical diet of the ruminant. Future insights into the microbial composition could be important to better understand the role of microbiotas in cattle animals (Figure 4).

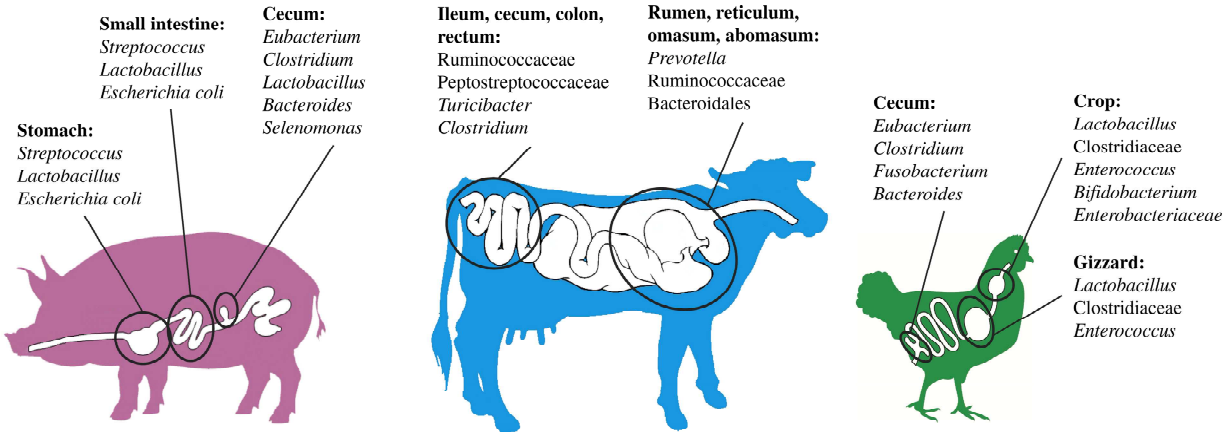


Figure 4: Overview of the gut microbiota composition of farm animals, i.e. pig, bovine and chicken.

Chickens microbiota

Chickens are considered an efficient agricultural species in converting feed to lean meat. Recent studies confirmed the assumption that the gut microbiota of a healthy chicken play (Figure 4) an important role in nutrient assimilation, vitamin and amino acid production and prevention of pathogen colonization (Cisek & Binek, 2014). Moreover, the chicken gut microbiota might also act as a source of bacterial pathogens that can expand to human beings, or act as a reservoir of antibiotic resistance determinants, which can be transferred to other microorganisms including opportunistic pathogens (Zhou*, et al.*, 2012). In poultry livestock, antibiotics are usually used to enhance growth performance, but they are also employed as a prophylactic therapy to prevent disease development and transmission (Allen & Stanton, 2014). On the other hand, antibiotic therapy is known to change the microbial composition, thus catalyzing dysbiosis with a resulting disadvantageous impact on physiology and metabolic performance of the host that may eventually result in the development of gut disorders (Allen & Stanton, 2014). Consequently, a greater number of studies concern the chicken microbiota. The chicken gut is subdivided in three upper segments, encompassing the crop, the proventriculus and the gizzard. The crop is a food storage muscular pouch involved in the breakdown of starch and in the fermentation of lactate. In the proventriculus start the digestion of the diet, while the gizzard mills food. The gizzard presents a lower pH and fermentation activity and acts as microbial barrier. Similar microbial communities were found in the crop and gizzard. The most dominant bacteria present in these two segments are lactobacilli, facultative and microaerophilic bacteria. Other abundant species belonged to Clostridiaceae, *Enterococcus* and in the case of the crop also *Bifidobacterium* and Enterobacteriaceae (Figure 4) (Deusch*, et al.*, 2015).

The small intestine consists of three parts, involving the duodenum, the jejunum and the ileum where the nutrient absorption and food digestion occurs. Due to the low pH, pancreatic and bile secretions, the bacterial density in the duodenum is comparably low. Notably, *Lactobacillus* and *Streptococcus* are the main colonizer of the jejunum. The chicken's ileum harbors higher abundance of bacteria

belonging to the genus *Lactobacillus* and in lower amount *Streptococcus*, Enterobacteriaceae and Clostridiaceae (van der Wielen*, et al.*, 2002, Gong*, et al.*, 2007). The ileum is also known to be colonized by butyrate producing bacteria that may play a pivotal role regarding the availability of nutrients, absorption rate and chicken performance

The most complex microbial community within the chicken gut is the one resident in the cecum (Figure 4), which has been shown to be dominated by the phyla Firmicutes, Bacteroidetes, Actinobacteria and Proteobacteria (Sergeant*, et al.*, 2014). Chickens have two caeca which are important for recycling urea, the absorption of water, and digestion of cellulose, starch and polysaccharides (Deusch*, et al.*, 2015). Fermentation of starch polysaccharides leads to the production of short chain fatty acids (SCFAs) that are adsorbed and catabolized by the host, contributing to animal nutrition and to inhibition of acid-sensitive pathogens (Turnbaugh*, et al.*, 2006). Previously studies reported a positive correlation between poultry cecal microbiota and the efficiency of energy extraction from feed (Stanley*, et al.*, 2013). Furthermore, a number of bacterial phylotypes were identified that significantly differed in richness between birds with high and low apparent metabolizable energy extraction ability (Stanley*, et al.*, 2013).

D. Gut microbiota and microbial biomarkers

Several recent metagenomic studies regarding gut microbiota allowed to identify specific alterations of bacterial composition that are correlated to multifactorial causes, such as diet, geography and health state. These scientific evidences support microbiota composition as an emerging potential biomarker or indicator of health. Identification of specific microbial biomarkers might help predict risk of chronic diseases such as metabolic syndrome, colon cancer and intestinal inflammations. Furthermore, microbiota interventions, i.e. bacterial therapies, might offer a new opportunity for targeting chronic diseases in the future. Increase knowledge about the interaction between commensal microbiota, host immunity and host metabolism may result in better understanding of diseases, even beyond the gastrointestinal tract and lead to better preventive and therapeutic strategies

Table 1 describes the main microbial biomarkers associated with various diseases that have be so far proposed.

Below is reported a short summary of the main microbial biomarkers so far identified through the comparison between healthy and disease gut microbiota. In detail, analysis on CRC individuals displayed instability in the composition of their gut bacterial communities when compared with healthy controls (Castellarin*, et al.*, 2012, Liang*, et al.*, 2017). An enrichment of *Fusobacterium nucleatum* and *Bacteroides fragilis* has been shown in CRC samples (Wei*, et al.*, 2016) and inversely correlated to the CD3+ T cell density, providing mechanistic evidence for the interactive roles of this microorganism in adaptive immunity (Mima*, et al.*, 2015).

Moreover IBD, including the major IBD phenotypes ulcerative colitis (UC) and Crohn's disease (CD), have been associated with higher abundance of taxa such as Enterobacteriaceae, *Ruminococcus gnavus* and *Desulfovibrio*, as well as less abundance of as *Faecalibacterium prausnitzii*, Lachnospiraceae and *Akkermansia*, compared to healthy condition (Berry & Reinisch, 2013).

Investigation of microbiota composition of individuals affected by *Clostridium difficile* (CDI) indicated an association with putative protective components of the microbiota, like *Alistipes*, *Butyricimonas* and *Oscillospira* that were significantly underrepresented in CDI samples (Milani*, et al.*, 2016). On the other side, a large number of opportunistic pathogens were overrepresented in faecal samples collected from individuals affected by CDI (Milani*, et al.*, 2016).

Additionally, recent study identified a specific bacterial family, i.e. Christensenellaceae, that appeared to be associated with a low body mass index (BMI), or leanness. In detail, germ-free mice that were treated with *Christensenella minuta* gained less weight than untreated mice, suggesting that increasing the amounts of this microbe may help to prevent or reduce obesity (Ley*, et al.*, 2006).

Table 1. List of intestinal disease biomarkers identified in public studies

Disease	Articles	Higher in Disease		Higher in Healthy	
		Phylum	**Genera**	**Phylum**	**Genera**
Crohn's Disease	PMID: 26313691	Firmicutes	*Anaerostipes*	Firmicutes	*Christensenella*
			Coprococcus		*Clostridium*
			Lachnobacterium		*Anaerofustis*
			Veillonella		*Lachnospira*
			Eubacterium		*Roseburia*
					Faecalibacterium
					Ruminococcus
		Bacteroidetes	*Odoribacter*	Bacteroidetes	*Butyricimonas*
					Alistipes
					Bacteroides
		Proteobacteria	*Escherichia/Shigella*		
			Proteus		
	PMID: 26922889			Firmicutes	*Faecalibacterium*
					Clostridium
				Bacteroidetes	*Prevotella*
		Proteobacteria	*Escherichia/Shigella*		
			Pseudomonas		
			Morganella		
		Fusobacteria	*Fusobacterium*		
	PMID: 24629344	Firmicutes	*Veillonella*	Firmicutes	*Dialister*
					Lachospiraceae group
					Coprococcus
					Oscillospira
					Faecalibacterium
					Dorea
					Erysipelotrichaceae
					Ruminococcacea group
				Bacteroidetes	Rikenellaceae
					Parabacteroides
					Sutterella
		Proteobacteria	*Haemophilus*	Proteobacteria	*Bilophila*
			Escherichia/Shigella		
		Fusobacteria	*Fusobacterium*		

Colorectal Cancer	PMID: 23940645			Firmicutes	*Ruminococcus*
					Pseudobutyrivibrio
				Bacteroidetes	*Bacteroides*
					Prevotella
		Verrucomicrobia	*Akkermansia*		
	PMID: 25104642	Firmicutes	Lachnospiraceae group	Firmicutes	Clostridiales group
					Lachnospiraceae group
					Staphylococcous
					Ruminococcaceae group
		Bacteroidetes	*Porphyromonas*	Bacteroidetes	*Bacteroides*
		Proteobacteria	Enterobacteriaceae group		
		Fusobacteria	*Fusobacterium*		
	PMID: 22009990	Fusobacteria	*Fusobacterium*		
		Firmicutes	Streptococcaceae group		
	PMID: 23497613	Firmicutes	*Roseburia*		
	PMID: 23733170	Firmicutes	Eubacteriaceae group	Firmicutes	*Faecalibacterium*
			Clostridiales group		*Roseburia*
				Bacteroidetes	Porphymonadaceae group
		Fusobacteria	*Fusobacterium*		
		Proteobacteria	*Campylobacter*		
	PMID: 26170900			Firmicutes	*Faecalibacterium*
					Lachnospiraceae group
					Ruminococcaceae group
				Bacteroidetes	*Bacteroides (uniformis)*
					Rikenellaceae group
		Fusobacteria	*Fusobacterium*		
		Proteobacteria	*Providencia*		

Clostridium difficile infection	PMID: 26383014	Firmicutes	*Clostridium*	Firmicutes	Ruminococcaceae group
			Enterococcaceae group		Lachnospiraceae group
			Streptococcaceae group		*Blautia*
			Lactobacillaceae group		*Roseburia*
					Dorea
				Bacteroidetes	*Bacteroides*
					Alistipes
		Actinobacteria	Actinomycetales	Actinobacteria	*Collinsella*
			Coriobacteriales		
		Proteobacteria	Gammaproteobacteria group		
			Escherichia/shigella		
			Proteus		
			Klebsiella		
			Providencia		
		Fusobacteria	*Fusobacterium*		
	PMID: 27481036	Firmicutes	*Veillonella*		
			Streptococcus		
			Lachnospiraceae group		
		Bacteroidetes	*Parabacteroides*		
		Proteobacteria	*Enterobacteriaceae* group		
	PMID: 22699611			Bacteroidetes	*Bacteroides*
				Firmicutes	*Coprococcus*
					Roseburia
					Streptococcus
	PMID: 23804381	Firmicutes	Lachnospiraceae group	Firmicutes	*Faecalibacterium*
			Peptostreptococcaceae		*Streptococcus*
			Enterococcus		*Ruminococcus*
			Lactobacillus		*Dorea*
					Lachnospiraceae group
					Blautia
					Roseburia
					Pseudobutyvibrio
					Anaerostipes
					Subdoligranulum
		Verrucomicrobia	*Akkermansia*		
		Bacteroidetes	*Alistipes*		
			Parabacteroides		
	PMID: 27166072	Actinobacteria	*Corynebacterium*	Actinobacteria	*Bifidobacterium*
				Bacteroidetes	*Bacteroides*
					Alistipes
					Barnesiella
		Firmicutes	Peptostreptococcaceae	Firmicutes	*Lachnospira*
			Enterococcus		*Faecalibacterium*
			Veillonella		Ruminococcaceae group
			Staphylococcus		Lachnospiraceae group
		Proteobacteria	*Klebsiella*		
			Escherichia/Shigella		
			Sutterella		
			Citrobacter		
			Proteus		
			Helicobacter		
			Morganella		
			Hafnia		
				Verrucomicrobia	*Akkermansia*
	PMID: 25756679	Firmicutes	*Phascolarctobacterium*	Firmicutes	*Blautia*
			Enterococcus		*Holdemania*
			Clostridium XI		Veillonellaceae group
			Flavinofractor		
			Erysipelotrichaceae		
				Proteobacteria	Enterobacteriaceae group
	PMID: 27531910	Firmicutes	*Streptococcus*		

Ulcerative Colitis			*Enterococcus*		
		Actinobacteria	*Bifidobacterium*		
	PMID: 27217061			Firmicutes	*Roseburia*
				Verrucomicrobia	*Akkermansia*
		Proteobacteria	*Haemophilus*		
			Phenylobacterium		
	PMID: 27604252			Actinobacteria	*Collinsella*
				Bacteroidetes	*Coprobacter*
					Paraprevotella
					Odoribacter
				Firmicutes	vadinBB60 group
					Christensenellaceae group
					Anaerotruncus
					Ruminococcaceae group
					Peptostreptococcaceae group
				Proteobacteria	*Thalassospira*
					Parasutterella

Chapter 2

Outline of the thesis

The focus of this Ph.D. thesis is to characterize the gut microbiota of animal and human beings. In fact, it is now well established that a healthy gut microbiota is largely responsible for overall health of the host. The alteration of the normal gut microbiota, i.e. dysbiosis, have been associated with a large array of human and animal diseases. Depth metagenomics analysis, such as 16S rRNA profiling and shotgun metagenomics, of the gut microbiota of animal and human datasets allowed to collect information that permitted to better understand the host-microbial interactions.

Chapter 3 shows a comparison of the gut microbiota of chickens used for large-scale commercial production and those grown in semi-wild conditions in order to investigate compositional differences of the cecal microbiota of poultry kept under different housing regimes and nutritional circumstances.

Chapter 4 illustrates an in-depth cataloguing of the microbiota composition of free-range, broiler and feral chickens by means of omics approaches that used metagenomic and culturomics techniques. Furthermore, various bacterial cultivation approaches based on multiple culture conditions were applied.

Chapter 5 describes meta-analysis of the human gut microbiome with aim to confirm the notion that urbanization/industrialization processes have substantially influenced the composition and functionality of the human gut microbiome.

Chapter 6 discusses the role of the intestinal bacteria in development of gastrointestinal disorder. In particular, a metagenomics analysis has been performed with the aims to identify a statistically significant and comprehensive correlation between microbiome and functional constipation.

Chapter 7 describes a meta-analysis evaluating the accuracy of different 16S rRNA gene-targeting PCR primers, and evaluate their impact on the profiling of the gut microbiota. Furthermore, potential universal intestinal disease microbial biomarkers were identified through cross-disease comparisons.

Chapter 3

Insights into the biodiversity of the gut microbiota of broiler chickens

Mancabelli L, Ferrario C, Milani C, Mangifesta M, Turroni F, Duranti S, Lugli GA, Viappiani A, Ossiprandi MC, van Sinderen D, Ventura M.

This chapter was published in:

Environmental Microbiology. 2016 Dec;18(12):4727-4738. Epub 2016 Jun 27.

environmental microbiology

Environmental Microbiology (2016) **18**(12), 4727–4738 doi:10.1111/1462-2920.13363

Insights into the biodiversity of the gut microbiota of broiler chickens

Leonardo Mancabelli,[1†] Chiara Ferrario,[1†] Christian Milani,[1] Marta Mangifesta,[2] Francesca Turroni,[1] Sabrina Duranti,[1] Gabriele Andrea Lugli,[1] Alice Viappiani,[2] Maria Cristina Ossiprandi,[3] Douwe van Sinderen[4] and Marco Ventura[1*]

[1]*Laboratory of Probiogenomics, Department of Life Sciences, University of Parma, Parco Area delle Scienze 11a, 43124 Parma, Italy.*

[2]*GenProbio srl, Parma, Italy.*

[3]*Department of Medical-Veterinary Science, University of Parma, Parma, Italy.*

[4]*APC Microbiome Institute and School of Microbiology, Bioscience Institute, National University of Ireland, Cork, Ireland.*

Summary

The gastrointestinal tract of poultry is densely populated with microorganisms, which are presumed to interact with the host and ingested feed. Comparison of the gut microbiota of chickens used for large-scale commercial production (Broiler Chicken, BC) and those grown in semi-wild conditions (Free-Range Chicken, FRC) revealed that at phylum level *Firmicutes* was the dominant phylum of the gut community in BC, while the gut microbiota of FRC contained higher levels of *Bacteroidetes* and *Proteobacteria*. Such differences may be due to the diet and/or the intensive use of antibiotics in BC. Indeed, analysis of the resistome of the cecal microbiomes showed a marked richness in BC datasets, with a modulation of the cecal microbiota toward antibiotic resistant bacteria. Functional characterization of the microbiome of FRC samples revealed an increase in gene pathways involved in degradation of complex carbohydrates. Furthermore, *in silico* analyses of the microbiomes of FRC and BC revealed a higher presence in genes involved in formate production in BC samples. Notably, compared to the BC microbiomes the FRC microbiomes were shown to contain a higher abundance of genes involved in the pathway for acetate production.

Received 1 February, 2016; revised 11 March, 2016; accepted 17 March, 2016. *For correspondence. E-mail marco.ventura@unipr.it; Tel. ++39-521-905666; Fax ++39-521-905604. †These authors contributed equally to this work.

Introduction

The bacterial taxa that constitute the large and complex microbial population resident in the gastrointestinal (GI) tract of chickens are believed to benefit the host by adding metabolic potential (Gerritsen et al., 2011), influencing host nutrition, gut development and physiology (Kau et al., 2011).

Chickens are considered to represent an efficient agricultural species in converting feed to lean meat, although their feed is often of low digestibility and their intestines are smaller, with shorter transit digestion times compared to those of mammals (Choct, 2009; McWhorter et al., 2009). The gut microbiota of a (healthy) chicken is presumed to play an important role in nutrient assimilation, vitamin and amino acid production and prevention of pathogen colonization (Apajalahti, 2005). The chicken gut microbiota may also act as a source of bacterial pathogens which can spread to human beings, or act as a reservoir of antibiotic-resistance determinants, which can be transferred to other microorganisms including opportunistic pathogens (Zhou et al., 2012). In poultry breeds, antibiotics are widely used to improve growth performance, while they are also employed as a prophylactic therapy so as to prevent disease development and transmission (Allen and Stanton, 2014). However, antibiotic therapy is known to cause substantial compositional alterations in microbial consortia, thus catalysing dysbiosis with a consequent detrimental impact on physiology and metabolic performance of the host that may ultimately result in the development of gut disorders (Allen and Stanton, 2014).

The most complex microbial community within the chicken gut is the one resident in the cecum, which has been shown to be dominated by the phyla *Firmicutes*, *Bacteroidetes*, *Actinobacteria* and *Proteobacteria* (Sergeant et al., 2014). In this context, the cecal microbiota of a chicken is implicated in nitrogen recycling from uric acid, producing essential amino acids and digestion of non-starch polysaccharides (NSPs) (Jozefiak et al., 2004).

Fermentation of NSPs leads to the production of short chain fatty acids (SCFAs) that are adsorbed and catabolized by the host, contributing substantially to animal nutrition and to inhibition of acid-sensitive pathogens (Turnbaugh et al., 2006). Previously, Stanley and colleagues (2013) reported the correlation between poultry cecal microbiota and the efficiency of energy extraction from feed. A number of bacterial phylotypes were identified that significantly differed in abundance between birds with high and low apparent metabolizable energy extraction ability.

Here, we performed an observational study directed to investigate the microbiota differences of 84 cecal samples collected from Broiler Chicken (BC) and Free-Range Chicken (FRC) animals. The aim of this study was to investigate compositional differences of the cecal microbiota of poultry kept under different housing regimes (either reared in a space-limiting, high-throughput production environment, or kept under semi-natural, roaming conditions, with or without antibiotic supplementation) and nutritional circumstances.

Results and discussion

16S rRNA profiling of FRC and BC

Cecal samples from 84 chickens were obtained in order to assess the microbiota composition of FRC and BC (respectively 35 and 49 samples) based on 16S rRNA-sequencing analysis as described previously (Milani et al., 2013). MiSeq-mediated sequencing of 84 samples produced a total of 6 335 983 sequencing reads with an average of 75 428 reads per sample (Table S1). Quality and chimera filtering produced a total of 5 508 392 filtered reads with an average of 65 576 filtered reads per sample, and ranging from 8230 to 189 725 reads (Table S1).

Assessment of rarefaction curves based on the Shannon and Chao1 biodiversity indexes calculated for 10 subsampling of sequenced read pools indicated that both curves tend to reach a plateau. Therefore, in all cases the obtained sequencing data was deemed adequate to cover the vast majority of biodiversity contained within the samples (Fig. 1A and B). Moreover, average rarefaction curves

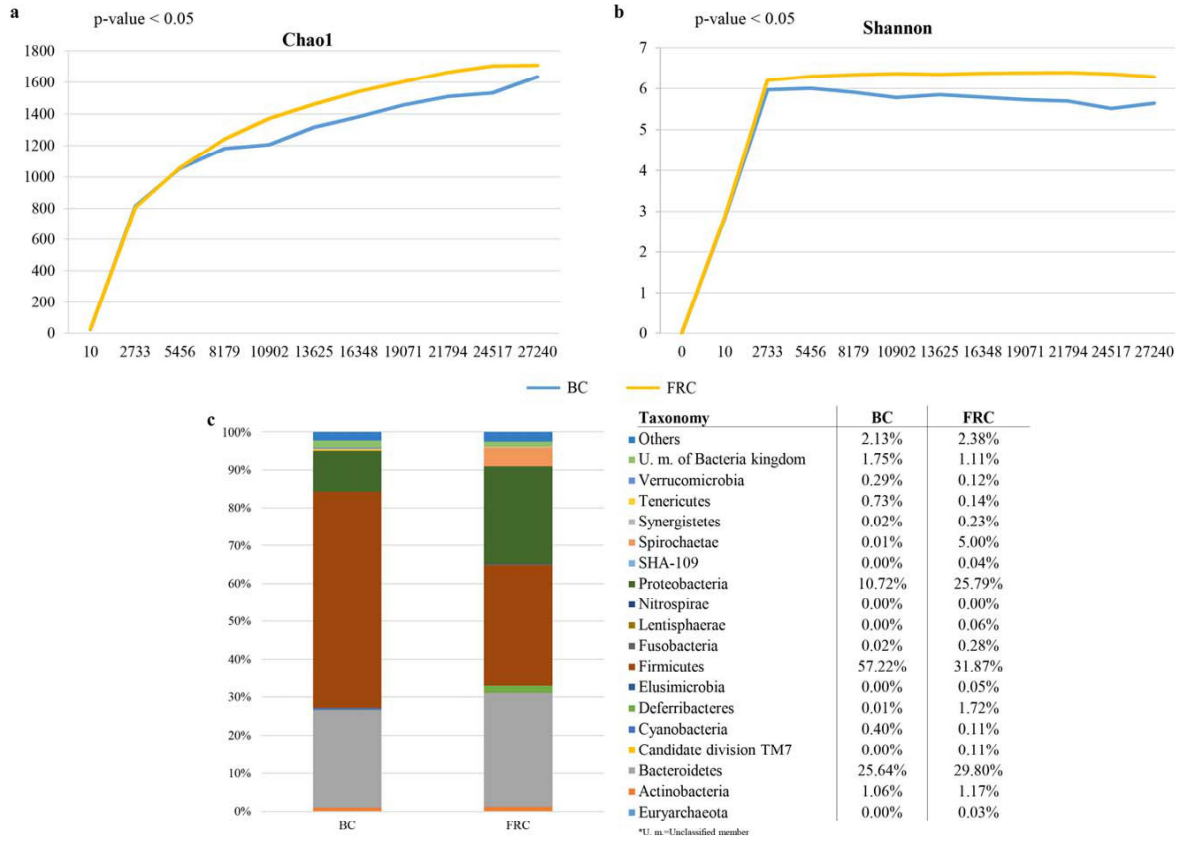


Fig. 1. Evaluation of alpha-diversity in BC and FRC samples. Panel A shows the average rarefaction curve representing variation of the Chao1 diversity index at increasing sequencing depth of BC and FRC samples. Panel B displays the average rarefaction curve representing variation of the Shannon diversity index at increasing sequencing depth of BC and FRC samples. Panel C represents a bar plot of the identified bacterial phyla in the 84 analysed samples. The legend reports the average of relative abundance of each phyla in both animal groups.

reveal a notable difference between the gut microbiota of FRC and BC poultry (Fig. 1). Specifically, cecal samples from FRC displayed a higher level of complexity of the gut microbiota compared to that found in BC samples (Fig. 1). The two curves are significantly different based on Student's t-test statistical analysis (t-value = 0.031, dfs = 68) calculated at the highest rarefaction depths reached by all the samples (Fig. 1).

Differential Gut microbiota composition

Inspection of predicted taxonomic profiles at phylum level for all samples exhibited that *Firmicutes* (58.90%) represented the dominant phylum of the cecal community in BC chickens, outnumbering the *Bacteroidetes* (25.70%) and *Proteobacteria* (10.73%) phyla. Differences in the gut microbiota composition were detected in FRC compared to BC with a comparative reduction of *Firmicutes* (31.86%) and a concomitant increase of *Bacteroidetes* and *Proteobacteria*, represented by 30.08% and 25.82% respectively (Fig. 1C). Studies in mice, pigs and humans have suggested that the *Firmicutes* to *Bacteroidetes* (F/B) ratio based on 16S rRNA gene sequence is often correlated with weight (Singh et al., 2013). The identified F/B ratio is 3.87 (P-value < 0.01) and 1.25 (P-value < 0.01) for BC and FRC groups (Fig. 1D) respectively. However, it is not possible to correlate performances and F/B ratio since no data about animal weight are available for the samples used in this study. Interestingly, antibiotic treatment has been shown to alter gut and fecal bacterial species composition in chicken (Gong et al., 2008) towards an increased abundance of *Lactobacillus* spp., Clostridiales and Enterobacteriaceae. Similar results were reported by Singh and colleagues (2013), who showed that an increase in the F/B ratio correlated with antibiotic treatment and increase of body weight, which may in turn be related to modulation of nutrient absorption by the host.

Furthermore, at a genus level, we identified 252 taxa of which 217 appear to be present in all samples, while six and 29 appeared to be uniquely present in cecal samples of BC and FRC respectively (Fig. 2A). Notably, the presence or absence of these genera may be the

Fig. 2. Exploration of the diversity in BC and FRC groups. Panel A depicts a Venn diagram illustrating the total, unique and shared number of OTUs predicted for BC and FRC datasets. Panel B reports the core set of genera present at an average relative abundance of $> 0.5\%$ in at least one sample group. Panel c shows the variation in terms of relative abundance $> 0.4\%$ and showing increase $> 50\%$ or decrease $< -30\%$ in FRC data sets as compared to those obtained from BC samples.

consequence of how the chickens are reared, i.e. being allowed to freely roam or being kept in a strictly confined environment, which is likely to impact on the physiology of the animals. In order to further identify differences in microbiota composition between BC and FRC groups, we focused on 97 genera, which were shown to be present at an average relative abundance of $> 0.5\%$ in at least one sample group (Fig. S1). The core set of genera obtained (Fig. 2B) and the comparison of the relative abundance average of BC and FRC (Fig. 2C) revealed a predominant presence of members of the *Firmicutes* phylum in BC samples, such as unclassified members (U. m.) of the Lachnospiraceae family, U. m. of the Ruminococcaceae family and *Intestinimonas* spp. In contrast, FRC-derived samples were shown to contain a higher diversity (compared to data obtained from BC samples) at the genus level among the *Bacteroidetes*, *Proteobacteria* and *Spirochaetes* phyla.

Notably, bacterial taxa belonging to both *Firmicutes* and *Bacteroidetes* phyla are known to be involved in the breakdown of otherwise indigestible (by the host) polysaccharides such as resistant starch and cellulose (Allen and Stanton, 2014; Stanley et al., 2012). Therefore, the observed differences in microbial composition between BC- or FRC-derived cecal samples may influence the food to energy conversion capacity in chickens housed under different conditions.

In order to evaluate microbiota differences between BC and FRC samples, we analysed the beta diversity based on unweighted UniFrac for these groups, after which the UniFrac distance matrix was represented through Principal Coordinate Analysis (PCoA) (Fig. 3). Interestingly, the samples were shown to group in two different clusters based on how the chickens had been kept, indicating that a different diet and/or prophylactic therapy impact on gut microbiota composition (Fig. 3A). Moreover, environmental samples included in this analysis, collected from BC and FRC litters, clustered with the corresponding cecal samples. Thus, the environment reflected the microbiota composition of chickens, perhaps due to contamination from the stool of the animals (Fig. 3A).

Evaluation of the beta diversity of cecal samples isolated from the BC and FRC groups, each following a particular feeding type (Fig. 3B), clearly shows a clear separation between BC and FRC animals. However, the microbiota composition of animals fed with different cereals, such as wheat and corn, clustered separately (Fig. 3B).

Analysis of the influence of antibiotic treatments on BC animals, showed a heterogeneous distribution of the samples (Fig. 3C), and it seems that different antibiotic mixes have a divergent impact on microbiota composition.

Such findings about PCoA clustering were confirmed by the obtained P-value of PERMANOVA statistical analysis (being < 0.05, dfs $= 1$), when the BC- and FRC-derived food supplementation and antibiotic treatment data sets are compared.

Prediction of the cecal microbiomes of FRC and BC

In order to evaluate the overall genetic content of the cecal microbiota of chickens, we determined the metagenome of 13 animals, representing five FRC and eight BC samples. Selection of these animals was based on the 16S rRNA microbial profiling data so as to include those birds that possess cecal microbial profiles that were closest to the average of their respective groups (Table 3).

NGS sequencing of these selected 13 samples (Table 3) produced a total of 3 379 001 raw reads that were filtered for human DNA and by quality, resulting in 331 216 filtered reads that were used for further analyses. When the taxonomic distribution predicted from the 16S rRNA profiling analysis was compared with that from the metagenomics data, there was a discrepancy in the ratio of *Firmicutes* and *Bacteroidetes*, probably due to differential amplification efficiency in the 16S rRNA profiling PCR, as previously reported (Sergeant et al., 2014). Nonetheless, the F/B ratio was shown to be higher in BC as compared to that obtained for FRC samples.

Moreover, functional classification of open reading frames based on the Cluster of Orthologous Genes (COG) obtained from assembled metagenomic datasets allowed detection of significant differences in relative abundance of COG functional categories between the two datasets. COG categories including replication, recombination and repair, energy production and conversion, carbohydrate transport and metabolism, as well as amino acid transport and metabolism were shown to be the most over-represented in both datasets (Fig. 4A). Interestingly, significant differences were detected for COG categories that are known to be involved in energy harvesting from food, such as energy production and conversion, carbohydrate transport and metabolism, amino acid transport and metabolism and lipid transport and metabolism categories (P-value < 0.05) (Fig. 4A and Table S2).

Identification of the resistome of the chicken gut microbiome

As above described, antibiotic therapy is widely used in breeding chickens (Pourabedin and Zhao, 2015). This practice impacts on the composition of the gut microbiota and is expected to influence their functionality (Allen and Stanton, 2014). However, very little is known about the occurrence of genes responsible for resistance against antibiotics, i.e. the resistome, in the chicken gut microbiome. In order to map and characterize the resistome of microbial consortia residing in the ceca of FRC and BC, the collective microbiomes of each of these two groups

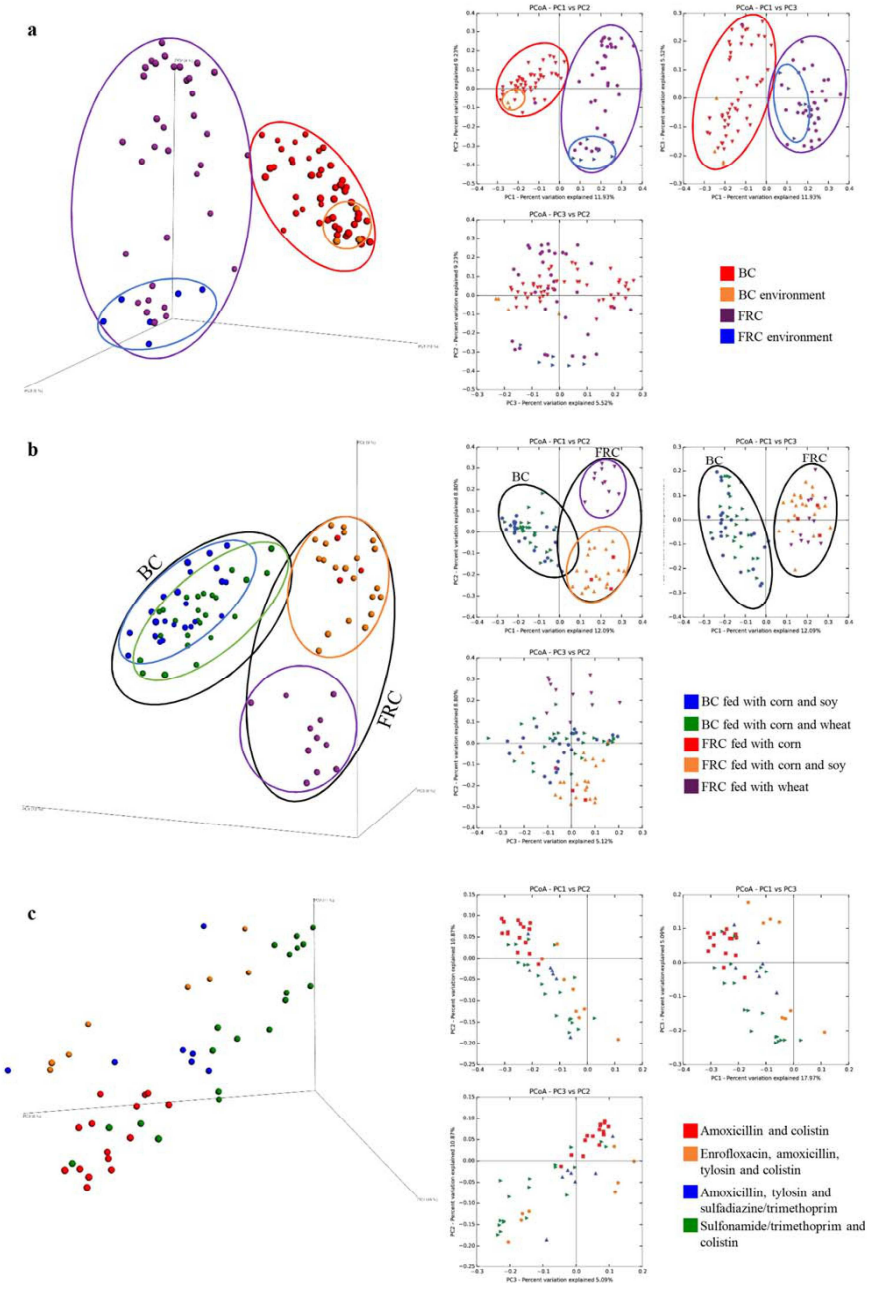


Fig. 3. Evaluation of the beta-diversity in BC and FRC samples. The predicted PCoA encompassing all 84 BC and FRC datasets is reported through two three-dimensional images as well as two-dimensional sections. Panel A shows BC, FRC and environmental datasets, and corresponding clusters are coloured in purple, red and blue/orange respectively. Panel B depicts the beta diversity of cecal samples from animals reared to different feeding types. Black circles represent BC and FRC groups, while coloured ones indicate the different cereal mix supplied. Panel c illustrates antibiotic treatments on BC animals. Antibiotic mix are reported in different colors.

were screened for known bacterial antibiotic-resistance genes (ARGs).

In silico analyses of shotgun metagenomic datasets revealed a higher abundance (16.64%) of ARGs in BC samples as compared to FRC animals (Fig. 4B). Interestingly, β-lactamase-encoding genes were shown to be the most abundant in all analyzed samples with a slight increase of 5.40% in BC samples. This might be the consequence of the fact that β-lactam-based antibiotics are the most commonly used drugs in BCs (Table 1), while β-lactamase-encoding genes are also the most commonly detected antibiotic resistance sequences found in soil and water bacteria (Allen and Stanton, 2014; Cho et al., 2014). Moreover, differences were noted in the heat map (Fig. 4B) regarding glycopeptide-based ARGs. In this context, we observed a 117.62% increase in the abundance of

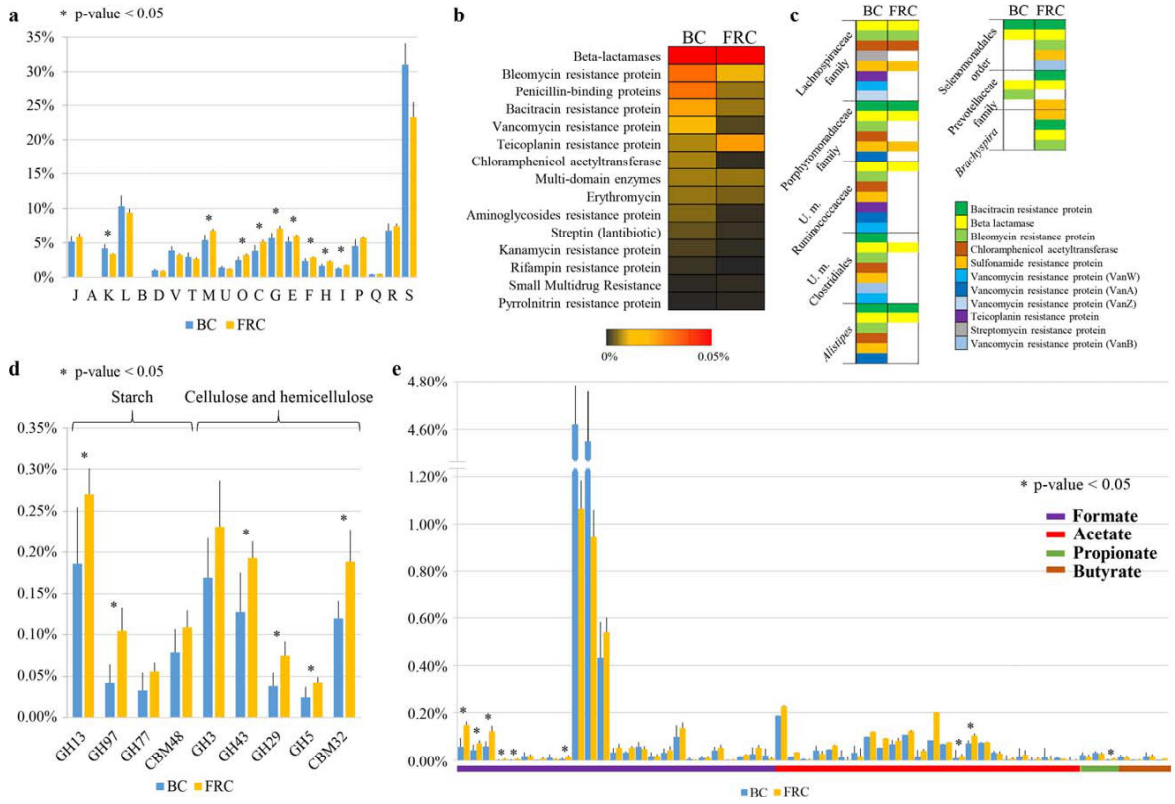


Fig. 4. Functional changes in the gut microbiome of BC and FRC. Asterisks indicate statistically significant differences between the two datasets (P-value < 0.05). Panel A depicts the functional annotation of BC and FRC metagenomic datasets according to COG categories. Each COG family is identified by a one-letter abbreviation (National Center for Biotechnology Information database). Panel B shows relative abundance of predicted enzymes involved in conveying antibiotic resistance as present in BC and FRC shotgun metagenomic datasets. Names of protein-coding genes are listed on the left, while names of sample groups used are listed at the top. Panel C exhibits ARGs in the bacterial taxa with higher variation in relative abundance in both datasets. Different colors represent various ARGs. Panel D shows changes in GH families involved in starch and plant cellulose/hemicellulose degradation in BC and FRC datasets. Panel E displays variation of pathways involved in formate, acetate, propionate and butyrate production in BC and FRC samples.

genes encoding for vancomycin resistance in BC samples. Conversely, we observed an increase (347.24%) of genes for resistance to teicoplanin in FRC poultry with respect to BC animals. Furthermore, an increase (9.15%) of chloramphenicol acetyltransferase-encoding genes were observed in BC samples.

Taxonomic classification of shotgun metagenomic reads corresponding to ARGs allowed the identification of bacterial taxa contributing to at least 1% of the total resistome (Fig. 4C). In the BC group, we observed a prevalent presence of ARGs belonging to the *Firmicutes* phylum (40.40%). In contrast, the FRC group contains a high abundance of *Bacteroidetes* (27.32%) and *Spirochaetes* (5.03%) phyla. Focusing on bacterial taxa that exhibit a higher variation in relative ARG abundance between BC and FRC samples, we detected an increased abundance of ARGs in BC taxa compared to FRC, as well as ARGs unique to the BC group encoding chloramphenicol acetyltransferase, teicoplanin resistance protein and vancomycin resistance protein (Fig. 4C).

Thus, our findings revealed a clear selection of the microbiota members resistant to these antibiotics. Furthermore, such data confirmed the positive correlation between supplementation of antibiotics and the F/B ratio increase as reported above (Fig. 1D).

Functional characterization of the chicken cecal microbiome

As mentioned above the reconstruction of the cecal microbiomes of chickens allowed their functional classification by means of the EggNog database (Powell et al., 2014). This analysis showed that a significantly higher number of functions associated with carbohydrate transport and metabolism are present in FRC datasets (7.1%) as compared to those obtained from BC (5.7%) (P-value < 0.05).

Table 1. Chicken samples collected in this study.

Sample name	Rearing methods	Origin	Feed	Antibiotics
P1	FRC	Parma 1 (Emilia Romagna, Italy)	Barley, wheat and wet waste	/
P2	FRC			
P3	FRC			
P4	FRC			
P5	FRC			
P6	FRC			
P7	FRC	Aulla (Toscana, Italy)	Wheat and wet waste	/
P8	FRC			
P9	FRC			
P10	BC	Cesena 1 (Emilia Romagna, Italy)	Cereals (wheat, corn), protein flour (soy, sunflower), vegetables oils (soy), mineral	Amoxicillin (20 mg Kg^{-1}) and colistin (50 mg Kg^{-1}).
P11	BC			
P12	BC			
P13	BC			
P14	BC			
P15	BC			
P16	BC			
P17	BC			
P18	BC			
P19	BC			
P20	BC	Cremona 1 (Lombardia, Italy)	Corn and soy	Amoxicillin (50 mg Kg^{-1}) and colistin (100 mg Kg^{-1}).
P21	BC			
P22	BC			
P23	BC			
P24	BC			
P25	FRC	Aulla (Toscana, Italy)	Wheat and wet waste	/
P26	FRC			
P27	BC	Cremona 2 (Lombardia, Italy)	Corn and soy	Sulfonamide/trimethoprim and colistin (100 mg Kg^{-1}).
P28	BC			
P29	BC			
P30	BC			
P31	BC			
P32	BC			
P34	BC			
P35	BC			
P36	BC			
P37	BC	Cremona 3 (Lombardia, Italy)	Corn and soy	Sulfonamide/trimethoprim and colistin (100 mg Kg^{-1}).
P38	BC			
P39	BC			
P40	BC			
P41	BC			
P42	BC			
P43	BC			
P44	BC			
P45	BC			
P46	BC			
P47	BC	Cesena 2 (Emilia Romagna, Italy)	Cereals (wheat, corn), protein flour (soy, sunflower), vegetables oils (soy), mineral	Enrofloxacin (10 mg Kg^{-1}), amoxicillin (20 mg Kg^{-1}), tylosin (50 mg Kg^{-1}) and colistin (50 mg Kg^{-1}).
P48	BC			
P49	BC			
P51	BC			
P53	BC			
P54	BC			
P55	BC			
P56	BC			
P57	BC	Cesena 3 (Emilia Romagna, Italy)	Cereals (wheat, corn), protein flour (soy, sunflower), animal fat (pig, cow, chicken oil) and minerals	Amoxicillin (20 mg Kg^{-1}), tylosin (50 mg Kg^{-1}) and sulfadiazine/trimethoprim (0.3 ml Kg^{-1}).
P59	BC			
P60	BC			
P61	BC			
P62	BC			
P63	BC			

Table 1. *cont.*

Sample name	Rearing methods	Origin	Feed	Antibiotics
P64	BC			
P68	FRC	Varese (Lombardia, Italy)	Wheat, soy and wet waste	/
P69	FRC			
P70	FRC	Reggio Emilia (Emilia Romagna, Italy)	Corn, soy, pea, bram and minerals	/
P71	FRC			
P72	FRC			
P73	FRC			
P74	FRC			
P75	FRC			
P76	FRC			
P77	FRC			
P78	FRC			
P79	FRC			
P80	FRC			
P81	FRC			
P82	FRC			
P83	FRC			
P84	FRC			
P85	FRC			
P86	FRC			
P87	FRC			
P88	FRC			
P89	FRC	Parma 2 (Emilia Romagna, Italy)	Corn, cereal mix and wet waste	/
P90	FRC			
P91	FRC			

Rearing methods, origin of samples, feed composition and antibiotic therapies are reported.

Moreover, we observed a difference in the COG family corresponding to energy production and conversion in FRC of 1.35% (P-value < 0.05) compared to BC datasets (Fig. 4A and Table S4). Complex polysaccharides are degraded by the gut microbiota into monosaccharides and then fermented to (mainly) produce the metabolic end products H_2, CO_2, CH_4 and SCFAs, which may then be adsorbed by the host together with any remaining digestible monosaccharides (Tremaroli and Backhed, 2012). The absorbed compounds promote hepatic triglyceride synthesis, the accumulation of triglycerides in adipocytes and subsequently, an increase in body fat (Janssen and Kersten, 2015).

In silico characterization of putative glycosyl hydrolases (GHs) responsible for degradation of complex carbohydrates, revealed that the microbiomes of FRC chickens possess a wider arsenal of GH families involved in starch, cellulose and hemicellulose degradation compared to BC samples (Fig. 4D). In particular, genes encoding predicted members of GH13, GH97 and GH77, as well as genes that specify proteins containing a CBM48 domain, are present at higher abundance (from 1.5- to 2-fold, P-value < 0.05) in the data sets from FRC animals compared to those of BC data (Fig. 4D and Table S5). GH genes encoded for α-amylases and α-glucosidases, while CBM48 is a carbohydrate-binding module known to bind various linear and cyclic α-glucans derived from starch and glycogen. Moreover, GH families GH3, GH43, GH29 and GH5, which represent α-L-arabinofuranosidase, β-D-xylopyranosidase and α-fucosidase activities (Matsuzawa et al., 2015), are more abundant in FRC datasets as compared to BC datasets (Fig. 4D). These differences can be explained by the higher abundancy of microorganisms belonging to the *Bacteroidetes* phylum in FRC chickens as compared to BC animals, where members of the former taxonomic groups are known to hydrolyze starch and plant cell wall compounds (e.g. cellulose, pectin and xylan) (Thomas et al., 2011).

Moreover, analysis of predicted bacterial metabolic pathways for SCFA production, showed a higher number of genes that are predicted to be responsible for formate production in BC compared to that found in FRC microbiomes (P-value < 0.05) (Fig. 4E, Tables S3 and S6). Formate produced in the animal gut can be degraded to CO_2 and H_2 (Sergeant et al., 2014). A high amount of hydrogen leads to reduction in fermentation and/or less energy-efficient fermentation to butyrate and propionate (Macfarlane and Macfarlane, 2003). The presence of bacteria that can metabolically act as a hydrogen sink, such as *Desulfovibrio*, *Helicobacter*, *Megamonas* and *Campylobacter* is expected to result in an increased efficiency of fermentation and SCFA production, which would ultimately exert a beneficial

effect to the host (Sergeant et al., 2014). These genera were present in the microbiota of both FRC and BC (Fig. S1).

In contrast, genes involved in acetate production were shown to be present at a significantly higher number (P-value < 0.05) in FRC microbiomes compared to BD datasets (Fig. 4E). This finding suggests a higher acetate production by the FRC microbiome, a notion that is supported by genes involved in acetate production, such as N-acetylglucosamine degradation or mycothiol biosynthesis pathways, which were shown to be significantly more abundant in FRC compared to BC (P-value < 0.05). Higher abundance of genes involved in acetate production in the microbiome of FRC animals may indirectly result in higher levels of butyrate through acetate-butyrate conversion, therefore confirming the results obtained by EggNog analysis.

In contrast, no difference in the abundance of genes predicted to be part of the metabolic pathways for butyrate and propionate synthesis was found between FRC and BC microbiomes with the exception of phosphatidylcholine resynthesis via the glycerophosphocholine pathway.

Conclusions

Improving growth performance in chickens has been one of the most important goals in poultry breeding. Recently, several studies have investigated the bacterial population that is resident in the chicken GI tract, using animals that had been reared under controlled conditions (Schokker et al., 2015; Stanley et al., 2013). The generated results suggest that the microbiota conveys benefits to the host by adding substantial metabolic potential to enhance nutrient utilization and energy conversion (Stanley et al., 2013; Waite and Taylor, 2015). Consistent with our results, these studies also found that the chicken cecal microbiota is dominated by *Firmicutes* and *Bacteroidetes*.

In this observational study, cecal samples from different slaughterhouses were collected and rearing conditions were listed to understand the microbiota composition of FRC and BC animals that had been kept at uncontrolled conditions.

The 16S rRNA gene microbial profiling data showed that the composition of the cecal microbiota of FRC is different from that of BC. In this context, *Firmicutes* dominate the BC animals, while FRC microbiota showed the predominance of bacteria belonging to the *Bacteroidetes* phylum. Furthermore, a distinct microbiota separation was detected between BC and FRC animals, apparently influenced by food supplementation and antibiotic treatment.

Analysis of the resistome of cecal microbiomes revealed that BC datasets contain higher levels of predicted ARGs compared to those of FRC, and a weak correlation was found with antibiotics used in the six different farms involved in this study (Table 1), particularly with amoxicillin. This observed increase in BC animals indicates that the use of antibiotics modulates the composition of the cecal microbiota toward antibiotic-resistant bacteria.

Functional characterization of the microbiome of FRC samples allowed the identification of key genetic features of the FRC microbiomes with respect to those of BC animals, such as an increase in the abundance of gene pathways involved in degradation of complex carbohydrates also encompassing those involved in most of the commercial chicken diets, grain. While for BC animals a higher abundance of genes involved in formate production was detected, in FRC data sets a higher number of genes was detected that are associated with acetate production, which in turn can be microbially converted to butyrate, a SCFA that can be adsorbed by the host as an energy source.

Experimental procedures

Ethic statement

In accordance with the REG CE No. 1099 of 2009 regarding the protection of animals at slaughter, this study did not require project license because no regulated procedures were carried out. Chickens were humanely killed at a designated establishment by cervical dislocation, which is an appropriate method recognized by REG CE No. 1099.

Animals and sampling

For the purpose of this study a total of 84 animals from two different poultry groups, i.e. BC (49 animals) and FRC (35 animals), were investigated (Table 1). BC were reared under commercial production conditions at six different farming centres from two geographical areas in Italy (Cesena and Cremona, Italy). These chickens were restricted in their movements due to high chicken density conditions and also they did not have access to foods other than that provided to them. All BC animals came from *Salmonella*-free breading. FRC were kept under semi-natural, free roaming conditions. This group is composed of animals from different geographical areas in the north of Italy (see Table 1). Feed composition and antibiotic treatments, i.e. name and dosage, are reported in Table 1.

All poultry GI tracts were recovered from different slaughterhouses where both BC and FRC animals were killed. Cecal samples were obtained, kept on ice and processed immediately after dissection. Briefly, each cecum was opened longitudinally. After removal of the digesta, 0.2 g of sample, composed of parts of both cecal pouches, were removed and briefly washed with RNA-later (Qiagen, Germany) to remove unattached or loosely attached bacteria from the walls. Samples were subjected to DNA extraction using the QIAamp DNA Stool Mini kit following the manufacturer's instructions (Qiagen).

Moreover, 10 environmental samples recovered from litters were included in this study. Samples were selected to represent the different rearing conditions associated with either large-scale, commercial production or the practice free-

Table 2. Environmental samples collected in this study.

Sample name	Origin	Sample type
E4	FRC – Parma 1 (Emilia Romagna, Italy)	FRC litter
E5	FRC – Parma 1 (Emilia Romagna, Italy)	FRC litter
E7	FRC – Aulla (Toscana, Italy)	FRC litter
E11	BC – Cesena 1 (Emilia Romagna, Italy)	BC litter
E14	BC – Cesena 1 (Emilia Romagna, Italy)	BC litter
E22	BC – Cremona 1 (Lombardia, Italy)	BC litter
E23	BC – Cremona 1 (Lombardia, Italy)	BC litter
E25	FRC – Aulla (Toscana, Italy)	FRC litter
E77	FRC – Reggio Emilia (Emilia Romagna, Italy)	FRC litter
E79	FRC – Reggio Emilia (Emilia Romagna, Italy)	FRC litter

Origin of samples are reported.

roaming (Table 2). DNA was extracted using the Power Viral environmental RNA/DNA kit (Mobio, USA) following the manufacturer's instructions.

16S rRNA gene amplification

Partial 16S rRNA gene sequences were amplified from extracted DNA using primer pair Probio_Uni and/Probio_Rev, which targets the V3 region of the 16S rRNA gene sequence (Milani et al., 2013). Illumina adapter overhang nucleotide sequences were added to the partial 16S rRNA gene-specific amplicons, which were further processed employing the 16S Metagenomic Sequencing Library Preparation Protocol (Part #15044223 Rev. B – Illumina; see also below). Amplifications were carried out using a Verity Thermocycler (Applied Biosystems). The integrity of the PCR amplicons were analyzed by electrophoresis on a 2200 TapeStation Instrument (Agilent Technologies, USA).

MiSeq sequencing of 16S rRNA gene-based amplicons

PCR products obtained following amplification of the 16S rRNA gene sequences were purified by magnetic purification step involving the Agencourt AMPure XP DNA purification beads (Beckman Coulter Genomics GmbH, Bernried, Germany) in order to remove primer dimers. DNA concentration of the amplified sequence library was determined by a fluorimetric Qubit quantification system (Life Technologies, USA). Amplicons were diluted to 4 nM and 5 μl of each diluted DNA amplicons were mixed to prepare the pooled final library. Sequencing was performed using an Illumina MiSeq sequencer with MiSeq Reagent Kit v3 chemicals.

16S rRNA-microbial profiling analysis

The fastq files were processed using QIIME (Caporaso et al., 2010) as previously described (Milani et al., 2013). Paired-end reads were merged and quality control retained sequences with a length between 140 and 400 bp, mean sequence quality score >25 and with truncation of a sequence at the first base if a low quality rolling 10 bp window was found. Sequences with mismatched forward and/or reverse primers were omitted.

In order to calculate downstream diversity measures (alpha and beta diversity indices, Unifrac analysis), 16S rRNA Operational Taxonomic Units (OTUs) were defined at ≥97% sequence homology using uclust (Edgar, 2010) and OTUs with less than 10 sequences were filtered. All reads were classified to the lowest possible taxonomic rank using QIIME (Caporaso et al., 2010) and a reference dataset from the SILVA database (Quast et al., 2013). Biodiversity of the samples (alpha-diversity) were calculated with Chao1 and Shannon indexes. Similarities between samples (beta-

Table 3. Shotgun metagenomic data.

Sample name	Origin	Number of reads	Final read number
Poultry 1	FRC – Parma 1 (Emilia Romagna, Italy)	260147	29283
Poultry 4	FRC – Parma 1 (Emilia Romagna, Italy)	263699	22556
Poultry 5	FRC – Parma 1 (Emilia Romagna, Italy)	229129	14840
Poultry 7	FRC – Aulla (Toscana, Italy)	233002	14355
Poultry 25	FRC – Aulla (Toscana, Italy)	216209	27550
Poultry 11	BC – Cesena 1 (Emilia Romagna, Italy)	377079	74856
Poultry 12	BC – Cesena 1 (Emilia Romagna, Italy)	295779	14669
Poultry 13	BC – Cesena 1 (Emilia Romagna, Italy)	387503	75081
Poultry 14	BC – Cesena 1 (Emilia Romagna, Italy)	324878	32319
Poultry 22	BC – Cremona 1 (Lombardia, Italy)	185450	10757
Poultry 23	BC – Cremona 1 (Lombardia, Italy)	157180	2301
Poultry 28	BC – Cremona 2 (Lombardia, Italy)	269347	10527
Poultry 36	BC – Cremona 2 (Lombardia, Italy)	179599	2122

diversity) were calculated by unweighted uniFrac (Lozupone and Knight, 2005). The range of similarities was calculated between the values 0 and 1. PCoA representations of beta-diversity were performed using QIIME (Caporaso et al., 2010).

Shotgun metagenomics

DNA was fragmented to 550–650 bp using a BioRuptor machine (Diagenode, Belgium). Samples were prepared following the TruSeq Nano DNA Sample Preparation Guide (Part#15041110Rev.D). Sequencing was performed using an Illumina MiSeq sequencer with MiSeq Reagent Kit v3 chemicals.

Analysis of metagenomic datasets

The generated fastq files were filtered for reads with a quality score of < 25, for sequences of chicken DNA, as well as for reads < 80 bp. Bases were also removed from the end of the reads unless the average quality score in a window of 5 bp was > 25. Only paired data were further analysed. The revised gene/protein set was searched using evolutionary genealogy of genes: Non-supervised Orthologous Groups (eggNOG; http://eggnog.embl.de/version_4.0.beta/) databases. Interrogation of sequence reads for significant identity to known ARGs was performed using a custom script based on RapSearch2 software (Zhao et al., 2012), htseq-count (Anders et al., 2015) and the database CARD (McArthur et al., 2013), which encompasses amino acidic sequences of enzymes involved in antibiotic resistance. Reconstruction of GH profiles as well as bacterial metabolic pathways and evaluation of their abundance in the shotgun metagenomics datasets was performed using custom scripts based on RapSearch2 software (Zhao et al., 2012), htseq-count (Anders et al., 2015) and the CAZy database or the MetaCyc database (Caspi et al., 2012) respectively.

Statistical analyses

ANOVA and PERMANOVA analyses were performed with SPSS software (www.ibm.com/software/it/analytics/spss/).

Data deposition

Raw sequences of 16S rRNA gene profiling are accessible through SRA study accession number SRP064851. Shotgun metagenomics data are accessible through SRA study accession number SRP064850.

Acknowledgments

We thank Dr. Francesco Biguzzi, Dr. Paolo Montagna and Dr. Franco Patercoli for kindly provided cecal samples and all informations regarding animal diets and antibiotic treatments. We thank GenProbio srl and Ipam srl for financial support of the Laboratory of Probiogenomics. LM is supported by Fondazione Cariparma, Parma, Italy. DvS is a member of the APC Microbiome Institute funded by Science Foundation Ireland (SFI), through the Irish Government's National Development Plan (Grant No SFI/12/RC/2273).

References

Allen, H.K., and Stanton, T.B. (2014) Altered egos: antibiotic effects on food animal microbiomes. *Annu Rev Microbiol* **68:** 297–315.

Anders, S., Pyl, P.T., and Huber, W. (2015) HTSeq–a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31:** 166–169.

Apajalahti, J. (2005) Comparative gut microflora, metabolic challenges, and potential opportunities. *J Appl Poult Res* **14:** 444–453.

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.

Caspi, R., Altman, T., Dreher, K., Fulcher, C.A., Subhraveti, P., Keseler, I.M. *et al.* (2012) The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. *Nucleic Acids Res* **40:** D742–D753.

Cho, H., Uehara, T., and Bernhardt, T.G. (2014) Beta-lactam antibiotics induce a lethal malfunctioning of the bacterial cell wall synthesis machinery. *Cell* **159:** 1300–1311.

Choct, M. (2009) Managing gut health through nutrition. *Br Poult Sci* **50:** 9–15.

Edgar, R.C. (2010) Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26:** 2460–2461.

Gerritsen, J., Smidt, H., Rijkers, G.T., and de Vos, W.M. (2011) Intestinal microbiota in human health and disease: the impact of probiotics. *Genes Nutr* **6:** 209–240.

Gong, J., Yu, H., Liu, T., Gill, J.J., Chambers, J.R., Wheatcroft, R., and Sabour, P.M. (2008) Effects of zinc bacitracin, bird age and access to range on bacterial microbiota in the ileum and caeca of broiler chickens. *J Appl Microbiol* **104:** 1372–1382.

Janssen, A.W.F., and Kersten, S. (2015) The role of the gut microbiota in metabolic health. *FASEB J* **29:** 3111–3123.

Jozefiak, D., Rutkowski, A., and Martin, S.A. (2004) Carbohydrate fermentation in the avian ceca: a review. *Anim Feed Sci Technol* **113:** 1–15.

Kau, A.L., Ahern, P.P., Griffin, N.W., Goodman, A.L., and Gordon, J.I. (2011) Human nutrition, the gut microbiome and the immune system. *Nature* **474:** 327–336.

Lozupone, C., and Knight, R. (2005) UniFrac: a new phylogenetic method for comparing microbial communities. *Appl Environ Microbiol* **71:** 8228–8235.

McArthur, A.G., Waglechner, N., Nizam, F., Yan, A., Azad, M.A., Baylay, A.J. *et al.* (2013) The comprehensive antibiotic resistance database. *Antimicrob Agents Chemother* **57:** 3348–3357.

Macfarlane, S., and Macfarlane, G.T. (2003) Regulation of short-chain fatty acid production. *Proc Nutr Soc* **62:** 67–72.

McWhorter, T.J., Caviedes-Vidal, E., and Karasov, W.H. (2009) The integration of digestion and osmoregulation in the avian gut. *Biol Rev Camb Philos Soc* **84:** 533–565.

Matsuzawa, T., Kaneko, S., and Yaoi, K. (2015) Screening, identification, and characterization of a GH43 family beta-xylosidase/alpha-arabinofuranosidase from a compost microbial metagenome. *Appl Microbiol Biotechnol* **99:** 8943–8954.

Milani, C., Hevia, A., Foroni, E., Duranti, S., Turroni, F., Lugli, G.A. *et al.* (2013) Assessing the fecal microbiota: an

optimized ion torrent 16S rRNA gene-based analysis protocol. *PLoS One* **8:** e68739.
Pourabedin, M., and Zhao, X. (2015) Prebiotics and gut microbiota in chickens. *FEMS Microbiol Lett* **362:** fnv122.
Powell, S., Forslund, K., Szklarczyk, D., Trachana, K., Roth, A., Huerta-Cepas, J. *et al.* (2014) eggNOG v4.0: nested orthology inference across 3686 organisms. *Nucleic Acids Res* **42:** D231–D239.
Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P. *et al.* (2013) The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* **41:** D590–D596.
Schokker, D., Veninga, G., Vastenhouw, S.A., Bossers, A., de Bree, F.M., Kaal-Lansbergen, L.M. *et al.* (2015) Early life microbial colonization of the gut and intestinal development differ between genetically divergent broiler lines. *BMC Genomics* **16:** 418.
Sergeant, M.J., Constantinidou, C., Cogan, T.A., Bedford, M.R., Penn, C.W., and Pallen, M.J. (2014) Extensive microbial and functional diversity within the chicken cecal microbiome. *PLoS One* **9:** e91941.
Singh, P., Karimi, A., Devendra, K., Waldroup, P.W., Cho, K.K., and Kwon, Y.M. (2013) Influence of penicillin on microbial diversity of the cecal microbiota in broiler chickens. *Poult Sci* **92:** 272–276.
Stanley, D., Keyburn, A.L., Denman, S.E., and Moore, R.J. (2012) Changes in the caecal microflora of chickens following Clostridium perfringens challenge to induce necrotic enteritis. *Vet Microbiol* **159:** 155–162.
Stanley, D., Geier, M.S., Denman, S.E., Haring, V.R., Crowley, T.M., Hughes, R.J., and Moore, R.J. (2013) Identification of chicken intestinal microbiota correlated with the efficiency of energy extraction from feed. *Vet Microbiol* **164:** 85–92.
Thomas, F., Hehemann, J.H., Rebuffet, E., Czjzek, M., and Michel, G. (2011) Environmental and gut bacteroidetes: the food connection. *Front Microbiol* **2:** 93.
Tremaroli, V., and Backhed, F. (2012) Functional interactions between the gut microbiota and host metabolism. *Nature* **489:** 242–249.
Turnbaugh, P.J., Ley, R.E., Mahowald, M.A., Magrini, V., Mardis, E.R., and Gordon, J.I. (2006) An obesity-associated gut microbiome with increased capacity for energy harvest. *Nature* **444:** 1027–1031.
Waite, D.W., and Taylor, M.W. (2015) Exploring the avian gut microbiota: current trends and future directions. *Front Microbiol* **6:** 673.
Zhao, Y., Tang, H., and Ye, Y. (2012) RAPSearch2: a fast and memory-efficient protein similarity search tool for next-generation sequencing data. *Bioinformatics* **28:** 125–126.
Zhou, W., Wang, Y., and Lin, J. (2012) Functional cloning and characterization of antibiotic resistance genes from the chicken gut microbiome. *Appl Environ Microbiol* **78:** 3028–3032.

Supporting information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Fig. S1. 16S rRNA-microbial profiling of the 84 chicken samples. Only taxa with a relative abundance of $> 0.5\%$ are shown. Sample names, origin and rearing methods are explained in Table 1 and in the figure.
Table S1. 16S rRNA microbial 10 profiling data.
Table S2. Average and standard deviation (SD) values of 13 the COG functional categories between the two datasets.
Table S3. Average and standard deviation (SD) values of the 17 predicted bacterial metabolic pathways for SCFA production between the two datasets.
Table S4. Degrees of Freedom (DFs), F distribution and significance 21 values obtained through Univariate ANOVA of the COG functional categories between the two datasets. Only statistically significant values were reported.
Table S5. Degrees of Freedom (DFs), F distribution and significance 26 values obtained through Univariate ANOVA of the Glycosyl Hydrolases (GH) and Carbohydrate Binging Modules (CBM) between the two datasets. Only statistically significant values were reported.
Table S6. Degrees of Freedom (DFs), F distribution and significance 31 values obtained through Univariate ANOVA of the predicted bacterial metabolic pathways for SCFA production between the two datasets. Only statistically significant values were reported.

Chapter 4

Untangling the cecal microbiota of feral chickens by culturomic and metagenomic analyses

Ferrario C, Alessandri G, Mancabelli L, Gering E, Mangifesta M, Milani C, Lugli GA, Viappiani A, Duranti S, Turroni F, Ossiprandi MC, Hiyashi R, Mackie R, van Sinderen D, Ventura M.

This chapter was published in:

Environmental Microbiology. 2017 Oct 2. [Epub ahead of print]

environmental microbiology

Environmental Microbiology (2017) **00**(00), 00–00 doi:10.1111/1462-2920.13943

Untangling the cecal microbiota of feral chickens by culturomic and metagenomic analyses

Chiara Ferrario,[1†] Giulia Alessandri,[1†] Leonardo Mancabelli,[1†] Eben Gering,[2] Marta Mangifesta,[3] Christian Milani,[1] Gabriele Andrea Lugli,[1] Alice Viappiani,[3] Sabrina Duranti,[1] Francesca Turroni,[1,4] Maria Cristina Ossiprandi,[4,5] Ricardo Hiyashi,[6] Roderick Mackie,[7] Douwe van Sinderen[8] and Marco Ventura[1,4*]

[1]*Laboratory of Probiogenomics, Department of Chemical Sciences, Life Sciences, and Environmental Sustainability, University of Parma, Parma, Italy.*

[2]*Kellogg Biological Station, Michigan State University, Department of Zoology, 3700 East Gull Lake Road, Hickory Corners, MI 49060, USA.*

[3]*GenProbio srl, Parma, Italy.*

[4]*Microbiome Research Hub, University of Parma, Parma, Italy.*

[5]*Department of Medical-Veterinary Science, University of Parma, Parma, Italy.*

[6]*Department of Animal Sciences, University of Illinois, Urbana, IL 61801, USA.*

[7]*Carle R. Woese Institute for Genomic Biology, Division of Nutritional Sciences, University of Illinois, Urbana, IL 61801, USA.*

[8]*APC Microbiome Institute and School of Microbiology, Bioscience Institute, National University of Ireland, Cork, Ireland.*

Summary

Different factors may modulate the gut microbiota of animals. In any particular environment, diet, genetic factors and human influences can shape the bacterial communities residing in the gastrointestinal tract. Metagenomic approaches have significantly expanded our knowledge on microbiota dynamics inside hosts, yet cultivation and isolation of bacterial members of these complex ecosystems may still be necessary to fully understand interactions between bacterial communities and their host. A dual approach, involving culture-independent and -dependent techniques, was used here to decipher the microbiota communities that inhabit the gastro intestinal tract of free-range, broiler and feral chickens. *In silico* analysis revealed the presence of a core microbiota that is typical of those animals that live in different geographical areas and that have limited contact with humans. Anthropic influences guide the metabolic potential and the presence of antibiotic resistance genes of these different bacterial communities. Culturomics attempts, based on different cultivation conditions, were applied to reconstruct *in vitro* the microbiota of feral chickens. A unique strain collection representing members of the four major phyla of the poultry microbiota was assembled, including bacterial strains that are not typically retrieved from the chicken gut.

Received 2 August, 2017; revised 19 September, 2017; accepted 22 September, 2017. *For correspondence. E-mail marco.ventura@unipr.it; Tel. ++39-521-905666; Fax ++39-521-905604. †These authors contributed equally to this work.

Introduction

Analysing the microbiota composition of breeding animals has gained growing interest because this allows a prediction of the compositional structure and associated metabolites of such communities, which are believed to fundamentally impact on all aspects of host physiology (Hiergeist *et al.*, 2015). The compositional analysis of the gut microbiota has enjoyed rapid advances in recent years thanks to metagenomic approaches based on high throughput sequencing methods (Fraher *et al.*, 2012; Weinstock, 2012; Milani *et al.*, 2013). However, in order to obtain detailed mechanistic insights into the roles of gut microorganisms with regards to interactions with their host and/or other bacteria in their natural ecosystem, isolation and cultivation of individual members of the gut microbiota is required (Browne *et al.*, 2016; Lagier *et al.*, 2016).

A large proportion of the gut microbiota of animals, including that of humans, is represented by not yet cultured microorganisms, including bacteria that are highly sensitive to oxygen and/or highly nutritionally exigent, thus preventing cultivation by classical laboratory approaches (Lagier *et al.*, 2016). Most of the current knowledge concerning the gut microbiota is still restricted to humans (Clemente *et al.*, 2012; Lozupone *et al.*, 2012; Marchesi *et al.*, 2016),

although information on the gut microbiota of breeding-animals, including poultry, is increasing (Oakley *et al.*, 2014; Mao *et al.*, 2015; McCormack *et al.*, 2017). In this context, both culture-independent and cultivation-based approaches have revealed that the majority of the human gut microbiota consists of members of the phyla *Bacteroidetes* and *Firmicutes*, yet many additional players (i.e., other bacterial phyla present in the human gut microbiota, principally *Proteobacteria*, *Actinobacteria*, *Fusobacteria* and *Verrucomicrobia*), are part of this complex ecosystem (Human Microbiome Project Consortium, 2012; Lozupone *et al.*, 2012; Mirzaei and Maurice, 2017). Similar results were reported for farm-based animals, such as broiler chickens (Mohd Shaufi *et al.*, 2015; Mancabelli *et al.*, 2016). Nevertheless, and in contrast to human beings, our current knowledge on the composition of the gut microbiota of chickens is still rather underdeveloped, despite that fact that these animals represent an important food resource for humans, while also being a potential reservoir of food-borne pathogens (Oakley *et al.*, 2014).

Feralization is a process by which domesticated species have escaped their controlled environment and colonized new habitats. It is also called 'domestication in reverse' (Gering *et al.*, 2015; Johnsson *et al.*, 2016), as it involves the removal of direct anthropogenic control over natural and sexual selection regimes. Animals, such as poultry, subjected to a feralization process could be useful to elucidate the impact of human care on the development of microbiota. With this aim, we report here an in-depth cataloguing of the microbiota composition of free-range, broiler and feral (i.e., formerly domesticated, wild-living) chickens (abbreviated here as FRC, BC and FC respectively) by means of omics approaches that used metagenomic and culturomics techniques, and also involved bacterial cultivation approaches based on multiple culture conditions (Lagier *et al.*, 2015; Lagier *et al.*, 2016). The results allowed us to understand what happens to the cecal microbiota of domesticated chickens during the process of feralization, and to obtain a unique chicken cecal microbiota strain collection, which includes representatives of the four major phyla characteristic of this complex bacterial community.

Results and discussion

Microbial diversity of feral chicken

The microbiota composition of 15 cecal samples of chickens originating from Bermuda (Table 1) was assessed based on 16S rRNA amplicon sequencing as described previously (Milani *et al.*, 2013). These animals comprised wild animals or chickens reared under distinct circumstances, with variable degrees of human control or interference, called Bermuda Feral, Free-Range Bermuda Broilers and Free-Range Bermuda Layers, as described in the Experimental Procedures section. Illumina-based 16S rRNA microbial profiling produced a total of 855 136 sequencing reads with an average of filtered 53 177 reads per sample (Supporting Information Table 1). The microbiota analysis of Bermuda poultry (encompassing Bermuda Feral, Free-Range Bermuda Broilers and Free-Range Bermuda Layers) obtained here, was compared with that of 84 previously characterized BC (broiler chicken) and FRC (free-range chicken) samples (Table 1) (Mancabelli *et al.*, 2016), where collection of biological materials and protocols for DNA isolation and 16S rRNA gene microbial profiling had been performed in an identical manner to those used here (Mancabelli *et al.*, 2016).

Assessment of rarefaction curves based on the Shannon and Simpson biodiversity indexes calculated for ten subsamplings of sequenced read pools indicated that both curves tend to reach a plateau. Therefore, in all cases the obtained sequencing data was deemed adequate to cover the vast majority of biodiversity contained within the samples (Fig. 1A and B). Moreover, average rarefaction curves reveal a difference between the cecal microbiota of BC, FRC and Bermuda poultry (Fig. 1). Specifically, cecal samples from Bermuda displayed a higher level of complexity of the cecal microbiota compared to that found in BC samples (Fig. 1) (Student's *t*-test statistical analysis p-value < 0.05). In contrast, the curves are not significantly different between Bermuda poultry and FRC (p-value > 0.05) calculated at the highest rarefaction depths reached by all samples (Fig. 1).

The microbiota of Bermuda Feral was shown to be dominated by the *Bacteroidetes* phylum (average value 40%), followed by *Firmicutes* and *Proteobaceria* (at average values of, respectively, 30% and 18%). In particular, the most abundant bacterial taxa are Unclassified Member (U. m.) of the *Prevotellaceae* family (11.74%), *Bacteroides* spp. (9.13%) and U. m. of *Bacteroidales* order (5.46%) (Supporting Information Figures S1, S2 and S3). Non-significant differences in microbiota composition were observed in Free-Range Bermuda Broilers, and Free-Range Bermuda Layers compared to Bermuda Ferals, at phylum level, despite the finding that the *Proteobacteria* phylum was shown to be present at a lower relative abundance (about 10% in both groups) (Supporting Information Figure S1). Both Free-Range Bermuda Broilers, and Free-Range Bermuda Layers were dominated by the *Bacteroides* genus (respectively 18% and 30%) followed by U. m. of *Ruminococcaceae* family (9% and 7% respectively) (Supporting Information Figure S3).

Differences in cecal microbiota composition between BC, FRC and FC chickens

In order to evaluate microbiota differences between Bermuda poultry, BC and FRC samples, we analysed the β-diversity based on unweighted UniFrac for these groups,

Table 1. Chicken samples collected in this study.

Samples name	Rearing methods	Origin	Feed	Antibiotics	References
BER50, BER52, BER53, BER55, BER56, BER57A, BER57B, BER58, BERXX, BER	FC	St. George (Bermuda)	Wet waste[a]	/	This study
BER100, BER101, BER102, BER103, BER104	FRC	St. George (Bermuda)	Wheat and wet waste, soy-free, corn-free	/	
P1, P2, P3, P4, P5, P6	FRC	Parma 1 (Emilia Romagna, Italy)	Barley, wheat and wet waste	/	(Mancabelli *et al.*, 2016)
P7, P8, P9	FRC	Aulla (Toscana, Italy)	Wheat and wet waste	/	
P10, P11, P12, P13, P14, P15, P16, P17, P18, P19	BC	Cesena 1 (Emilia Romagna, Italy)	Cereals (wheat, corn), protein flour (soy, sunflower), vegetables oils (soy), mineral	Amoxicillin (20 mg kg^{-1}) and colistin (50 mg kg^{-1}).	
P20, P21, P22, P23, P24	BC	Cremona 1 (Lombardia, Italy)	Corn and soy	Amoxicillin (50 mg kg^{-1}) and colistin (100 mg kg^{-1}).	
P25, P26	FRC	Aulla (Toscana, Italy)	Wheat and wet waste	/	
P27, P28, P29, P30, P31, P32, P33, P34, P35, P36	BC	Cremona 2 (Lombardia, Italy)	Corn and soy	Sulfonamide/trimethoprim and colistin (100 mg kg^{-1}).	
P37, P38, P39, P40, P41, P42, P43, P44, P45, P46	BC	Cremona 3 (Lombardia, Italy)	Corn and soy	Sulfonamide/trimethoprim and colistin (100 mg kg^{-1}).	
P47, P48, P49, P51, P53, P54, P55, P56	BC	Cesena 2 (Emilia Romagna, Italy)	Cereals (wheat, corn), protein flour (soy, sunflower), vegetables oils (soy), mineral	Enrofloxacin (10 mg kg^{-1}), amoxicillin (20 mg kg^{-1}), tylosin (50 mg kg^{-1}) and colistin (50 mg kg^{-1}).	
P57, P59, P60, P61, P62, P63, P64	BC	Cesena 3 (Emilia Romagna, Italy)	Cereals (wheat, corn), protein flour (soy, sunflower), animal fat (pig, cow, chicken oil) and minerals	Amoxicillin (20 mg kg^{-1}), tylosin (50 mg kg^{-1}) and sulfadiazine/trimethoprim (0.3 ml kg^{-1}).	
P68, P59	FRC	Varese (Lombardia, Italy)	Wheat, soy and wet waste	/	
P70, P71, P72, P73, P74, P75, P76, P77, P78, P79, P80, P81, P82, P83, P84, P85, P86, P87, P88	FRC	Reggio Emilia (Emilia Romagna, Italy)	Corn, soy, pea, bram and minerals	/	
P89, P90, P91	FRC	Parma 2 (Emilia Romagna, Italy)	Corn, cereal mix and wet waste	/	

a. Wet waste: invertebrates, seeds, shoots, household and business garbage.
Origin of samples, antibiotic treatments and feed composition are reported.

after which the UniFrac distance matrix was represented through Principal Coordinate Analysis (PCoA) (Fig. 1C). Such analyses showed that the Bermuda samples grouped in the FRC group, being clearly separated from BC. Interestingly, a detailed analysis of the FRC group revealed a split in two different clusters that do not correlate with the geographical origin of the animals (Fig. 1D). In detail, the first cluster, called the FRC dominant group (FRC-group) was composed of FRC samples derived from different geographical areas and or livestock in Italy (Parma 2, Varese and Reggio Emilia), while the second one, indicated as the Feral Chicken-dominant group (FC-group) encompassed samples collected in Italy (Parma 1 and Aulla) and all those obtained from Bermuda (Bermuda Ferals, Free-range Bermuda Broilers and Free-Range Bermuda Layers), despite differences in diet and rearing methods (Table 1). Such findings were shown to be statistically validated by a p-value of < 0.001, as obtained by PERMANOVA analysis, when the data sets of the two clusters were compared. Thus, chickens from different lifestyles (FC vs. FRC), ancestries and rearing localities can exhibit striking similarities in core cecal microbiota.

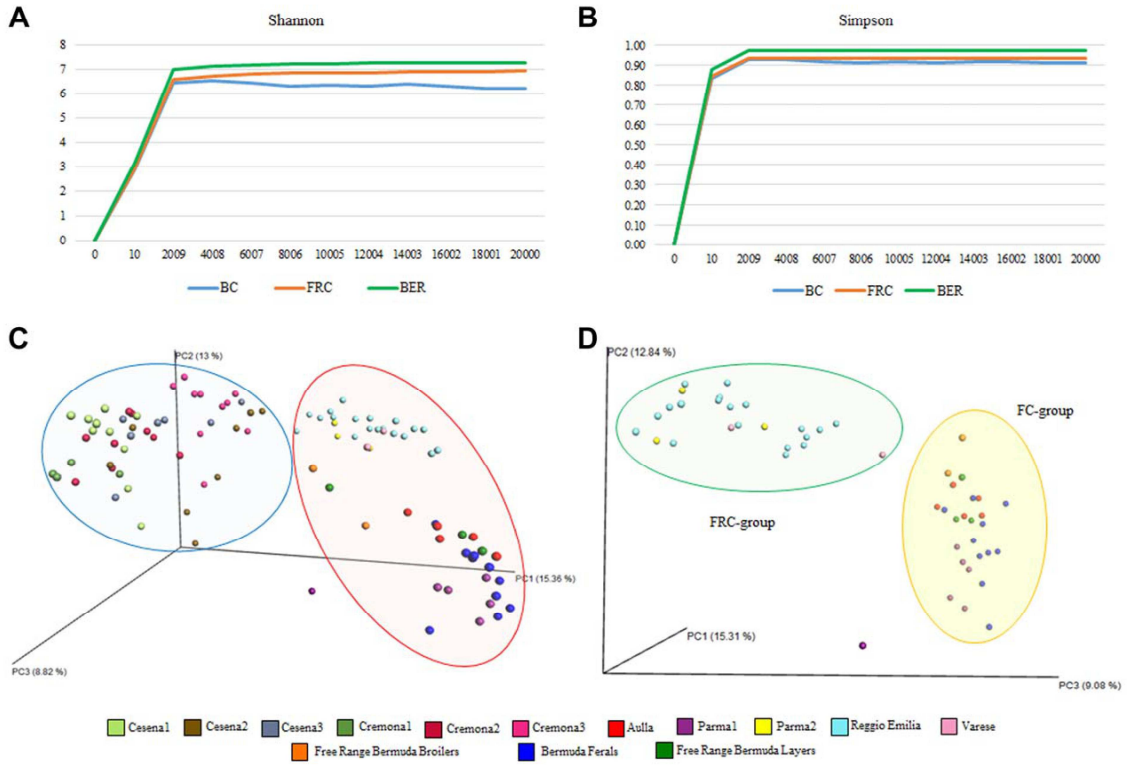


Fig. 1. Evaluation of α- and β- diversity in various poultry groups.
A. The average rarefaction curve representing variation of the Shannon diversity index at increasing sequencing depth of Bermuda, BC and FRC samples.
B. The average rarefaction curve representing variation of the Simpson diversity index at increasing sequencing depth of Bermuda, BC and FRC samples.
C. The predicted PCoA encompassing all Bermuda, BC and FRC data sets through three-dimensional images.
D. the predicted PCoA encompassing all Bermuda and FRC data sets through three-dimensional images.

Moreover, the existence of a core microbiota, being typical of chickens that live in different geographical areas with little if any human contact and feeding on a natural diet, was observed (Fig. 1C).

Inspection of the predicted microbiota in FRC, BC and FC groups

Exploring the predicted taxonomic profiles at phylum level for the analysed samples clearly shows that the BC, FRC and FC groups possess a distinct microbiota composition (Fig. 2A–C and Supporting Information Figs. S1, S2 and S3). Sample categorization in FRC, BC or FC was performed subsequent to obtaining the results from the PCoAs as reported in Fig. 1C and D. The bacterial community of the BC and FRC cecal samples was shown to be dominated by the *Firmicutes* phylum, although at a significantly different level (p-value < 0.05, with an average relative abundance of 59.8% and 33.4% respectively), outnumbering the *Bacteroidetes* (26.2% and 29.2%) and *Proteobacteria* (10.9% and 28.3%) phyla (Fig. 2A). Conversely, the FC-group cecal microbiota was dominated by members of the *Bacteroidetes* phylum (with an average relative abundance of 38.6%). *Firmicutes* in the FC-derived samples were present at a lower level in comparison with those from BC samples (32%, p-value < 0.05), while the *Proteobacteria* phylum was shown to be present at a significantly lower abundance (18.5%, p-value < 0.05) in comparison to the FRC cecal samples.

In order to further identify differences in microbiota composition between BC, FRC and FC, the microbiota composition was explored at genus level, evaluating differences of bacterial genera present at an average relative abundance of $> 0.5\%$ in at least one sample group (Fig. 2C and Supporting Information Fig. S3). The definition of the core microbiota, that is, bacterial species that are present in all samples of a given sample set (Salonen *et al.*, 2012), also revealed particular differences between FC, FRC and BC animals (Fig. 2D). In detail, unclassified members (U. m.) of the *Prevotellaceae* family, U. m. of the *Bacteroidales* order and *Succinimonas* spp. were preponderant in FC samples as compared to FRC and BC

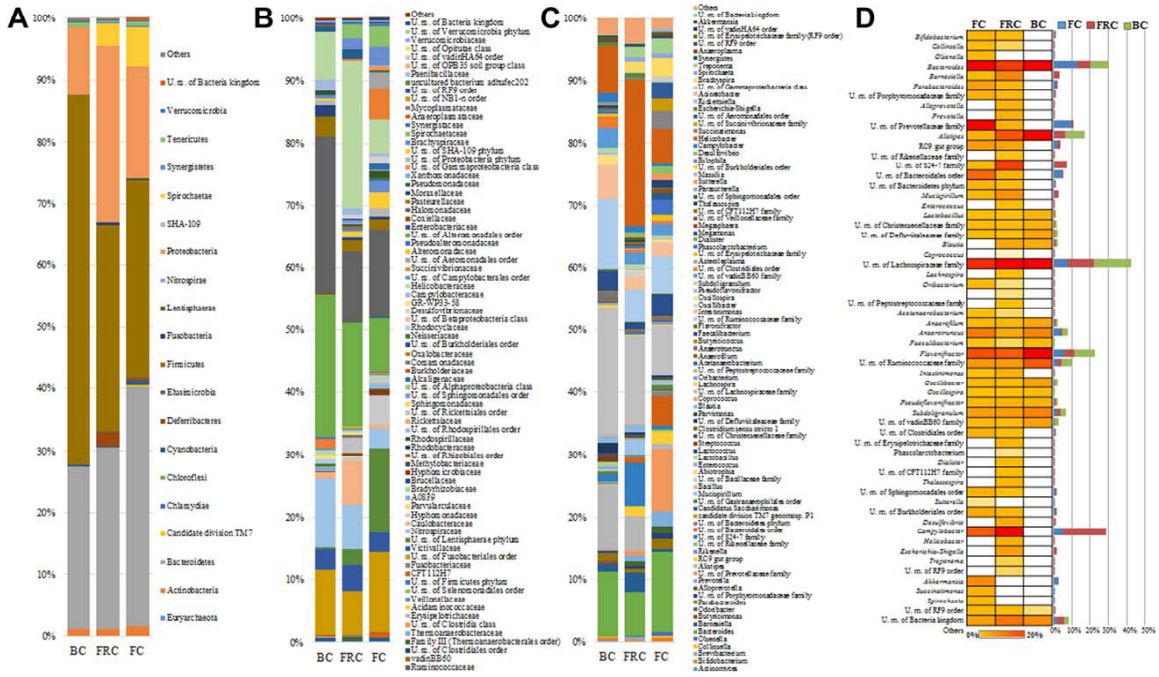


Fig. 2. Evaluation of the predicted microbiota in FRC, BC and FC groups. Sample categorization in FRC, BC or FC was performed subsequent to obtaining the PCoA results as reported in Fig. 1C and D.
A–C. The 16S rRNA-microbial profiling of aggregate FC, BC and FRC groups at phylum, family and genus level (only taxa with a relative abundance of $>0.5\%$ are shown) respectively.
D. A heat map reporting the presence/absence of core species detected in the microbiota of the three data sets. Moreover, the relevance of each genus in the FC, FRC and BC data sets is reported on the right. U. m.: unclassified member.

sample sets (p-value < 0.05) (Fig. 2C). In contrast, the cecal samples from the FRC-group showed a higher presence, at genus level, of U. m. of the *S24-7* family, *Barnesiella* spp., *Mucirospirillum* spp. and *Helicobacter* spp. The BC group elicited a preponderance in *Alistipes* spp., U.m. of *Ruminococcaceae* family and *Intestinimonas* spp. The microbial genera that were shared among the three chicken groups, representing a putative 'core' cecal microbiota, that is, defined as persistent members of a microbial community (Astudillo-Garcia *et al.*, 2017), which co-evolved with chickens in these different environments, were predicted to consist of 18 taxa (Fig. 2D). These core bacterial species belonged to the *Firmicutes* phylum, indicating the maintenance of a common core phylogeny under different ecological circumstances.

Prediction of the cecal microbiomes of FC

In order to evaluate the overall genetic content of the cecal microbiota of chickens, we decoded the microbiomes of three FC animals (BER, BER50 and BERXX, characterized as Bermuda Feral) by a metagenomics approach. Selection of these animals was based on the 16S rRNA microbial profiling data so as to include those birds showing cecal microbial profiles that were closest to the average of their corresponding group (Table 2). The obtained data were then compared to microbiome reconstruction information from a similar shotgun sequencing effort of three BC animals (Mancabelli *et al.*, 2016), selected to understand how anthropometric influences could drive the microbiota composition and its capability. Next Generation Sequencing of these selected six FC- and BC-samples (Table 2) produced a total of 53 219 426 raw reads that were filtered for human and poultry DNA and by quality, resulting in 13 876 919 filtered reads that were used for further analyses.

Notably, the taxonomic prediction achieved by the metagenomic analyses was consistent with that obtained by 16S rRNA profiling analysis, highlighting a preponderance in the relative abundance of bacteria belonging to the *Bacteroidetes* and to *Firmicutes* phyla in FC and BC animals respectively (Supporting Information Table S2 and Supporting Information Figure S4).

Functional characterization of the feral chicken cecal microbiome

The reconstruction of the cecal microbiome of feral chickens allowed functional classification through the EggNog

Table 2. Shotgun metagenomics raw data of FC and BC samples.

Sample name	Rearing category	Pe reads pre-filtering	Pe reads post-filtering
BER	FC	13819331	3260366
BER50	FC	19697499	5233692
BERXX	FC	18613136	4853492
P11	BC	377079	178602
P13	BC	387503	221710
P14	BC	324878	129057

database (Huerta-Cepas *et al.*, 2016). Data were compared with BC data sets (Mancabelli *et al.*, 2016) and significant differences (p-value < 0.05) were identified for the EggNog families encompassing Transcription, Cell wall/membrane/envelope biogenesis, Carbohydrate transport and metabolism, Nucleotide transport and metabolism and Inorganic ion transport and metabolism (Fig. 3A and Supporting Information Fig. S3).

Metabolic capability to convert complex carbohydrates from diet into simpler glycans and to subsequently transform them into short chain fatty acids (SCFAs) is driven by carbohydrate degrading enzymes, such as glycosyl hydrolases (GHs), and by several enzymes belonging to the SCFA biosynthetic pathways (den Besten *et al.*, 2013). *In silico* characterization of putative GHs responsible for the degradation of oligo/polysaccharides highlighted that the microbiomes of domesticated chickens (BC) possess a larger GH repertoire, in particular associated with the degradation of complex glycans derived from plant and cereals, such as mannan and (arabino)xylan (p-value <0.05) (Fig. 3B and Supporting Information Table S4). As reported previously, captive birds fed with grain were predicted to share a microbiota with increased capability for carbohydrate metabolism (Waite and Taylor, 2014). In particular, genes encoding β-galactosidases, α-amylases and β-glucosidases, encompassing members of the GH2, GH13 and GH3 families, were the most abundant in cecal microbiomes of BC (Fig 3B). These BC microbiomes were shown to be enriched in bacteria belonging to the *Firmicutes* phylum, which are able to degrade otherwise indigestible carbohydrates including cellulose and starch (Stanley *et al.*, 2013).

The level and types of SCFAs that are generated in the ceca are largely influenced by the amount of starch introduced in the ceca (Roto *et al.*, 2015). It is worth mentioning that the pronounced presence of genes for SCFA synthesis is related to acetate and formate production for both BC and FC animals (Fig. 3C). In this context, the FC microbiomes showed a reduced occurrence of genes predicted to encode enzymes involved in the production of formate, acetate, yet also propionate and butyrate (respectively, 11, 2.5, 4.3 and 3.6 time less) (Fig. 3C) in comparison to BC microbiomes. Indeed, only genes involved in a small number of pathways related to acetate (e.g., chitin degradation to ethanol, vanillin biosynthesis and rhamnogalacturonan type I degradation pathways) and formate production (e.g., tetrahydrobiopterin biosynthesis I and II, estradiol biosynthesis I and II, tryptophan degradation to 2-amino-3-carboxymuconate semi-aldehyde and formaldehyde oxidation) were enriched (p-value < 0.05) in the microbiomes of FC, compared to those of BC (Fig. 3D and Supporting Information Table S5).

The prevalence of genes related to carbohydrate metabolism and energy recovery from food is higher in the microbiomes of BC animals than those from Bermuda. This may be explained by the fact that farm animals have undergone a genetic selection toward maximum increase in body weight in a short time and a concurrent maximal conversion of diet into body weight. Thus, one may argue that the cecal microbiota of BC has been selected toward an enrichment of those microorganisms that perform a very efficient energy recovery from the diet (Ley *et al.*, 2008). In contrast, the FC animals have reduced dimensions, are more agile and fast (Gering *et al.*, 2015). Thus, microbiomes with a lower prevalence of genes related to carbohydrate metabolism and energy recovery from food, may convey a selective advantage for balancing mobility and body size.

Prediction of the resistome of feral chicken

In recent years increasing interest has emerged on how the gut microbiota may act as a reservoir for antibiotic resistance genes (ARGs) (Salyers *et al.*, 2004; Hu *et al.*, 2013; Yassour *et al.*, 2016). In this context, the investigation of ARGs in the poultry microbiome may represent a serious global health safety issue since such ARGs may be subject to horizontal transmission to pathogenic bacteria in the gut environment (von Wintersdorff *et al.*, 2016). For this reason, the FC resistome, that is, all genes predicted to be involved in antibiotic resistance (Wright, 2007), were mapped and compared with data previously obtained from BC animals (Mancabelli *et al.*, 2016), in which an antibiotic therapy was routinely administered (Table 1). *In silico* analysis of shotgun data revealed a higher abundance (70%) of ARGs in BC compared to FC animals. Interestingly, and as reported previously (Mancabelli *et al.*, 2016),

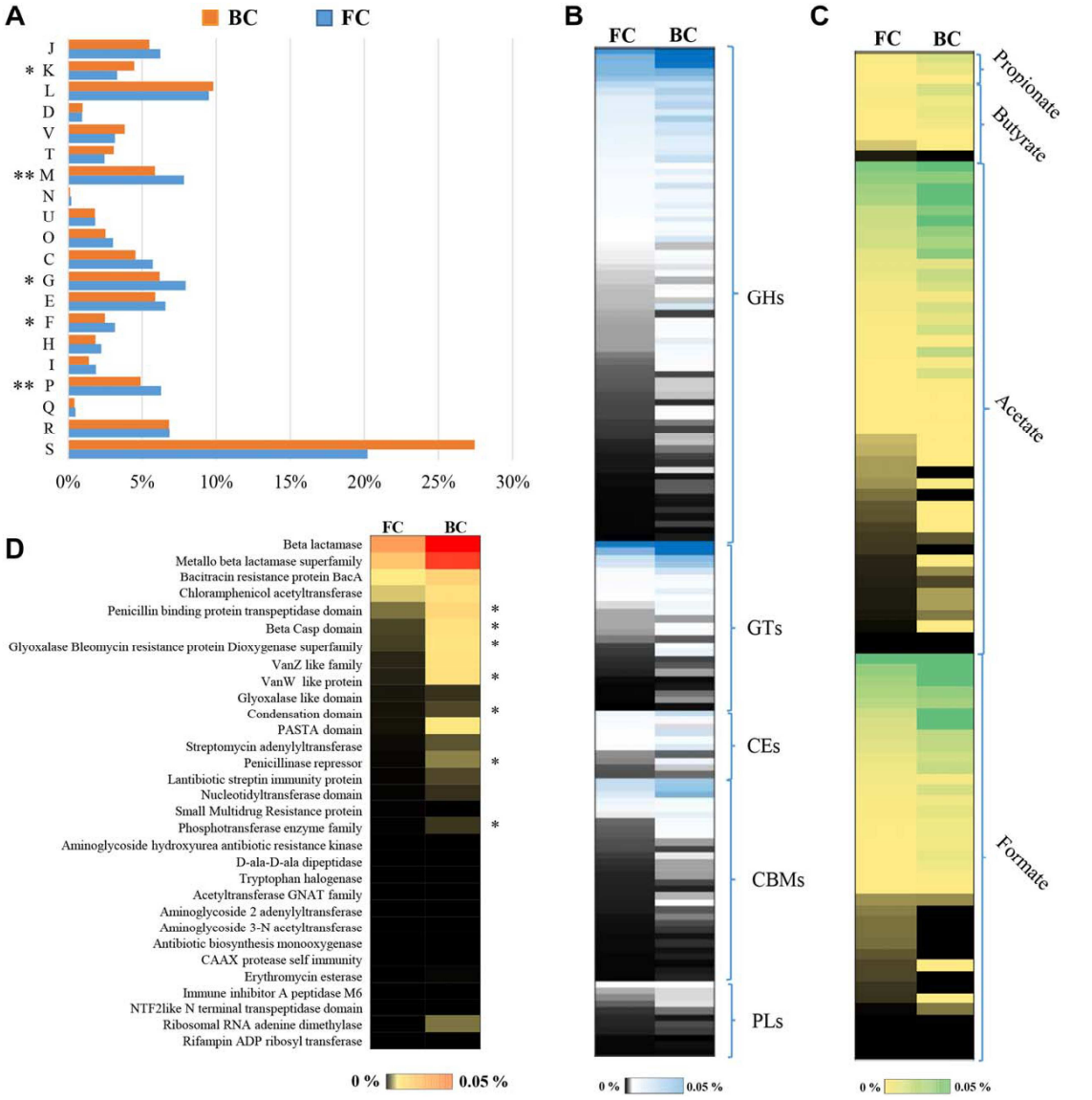


Fig. 3. Functional changes in the cecal microbiome of BC and FC. Asterisks indicate statistically significant differences between the two data sets (p-value < 0.05).
A. The functional annotation of BC and FRC metagenomic data sets according to COG categories. Each COG family is identified by a one-letter abbreviation (National Center for Biotechnology Information database).
B. Changes in carbohydrate degradation related genes in BC and FC data sets. GH, GT, CE, CBM and PL indicate, respectively, glycosyl hydrolase, glycosyl transferase, carbohydrate esterase, carbohydrate-binding module and polysaccharide lyase.
C. A heat map reporting variation of pathways involved in formate, acetate, propionate and butyrate production in BC and FRC samples.
D. Relative abundance of predicted enzymes involved in conveying antibiotic resistance as present in BC and FC shotgun metagenomic data sets. Names of protein-encoding genes are listed on the left, while names of sample groups used are listed at the top.

β-lactamase-encoding genes (as well as β-lactamase and metallo-β-lactamase superfamily) were the most abundant ARGs in all analysed samples (Fig. 3D), in particular in BC animals, which all had received amoxicillin, a β-lactamic antibiotic (Table 1).

Notably, the resistome of FC animals are enriched in ARGs encoding predicted β-lactamase activity and resistance to bacitracin or chloramphenicol (Fig. 3D and Supporting Information Table S6). The presence of ARGs genes in wild animals was previously demonstrated to be

related to human contact (Osterblad *et al.*, 2001; Thaller *et al.*, 2010). In this context, the feralization process with accompanying detachment from human care is correlated with the reduction of ARGs in the FC microbiome (Fig. 3D). Notably, a statistically significant difference (p-value < 0.05) was detected for genes related to β-lactamase, and resistance to bleomycin or vancomycin, with an increase of about 66%, 81% and 91% in BC samples respectively (Supporting Information Table S6).

In vitro *reconstruction of the core microbiota of FC poultry*

Culturomics approaches were applied to those cecal samples of FC and FRC animals that had been included in the FC-group (Fig. 1C and D) from the 16S rRNA gene microbial profiling. These attempts were developed in order to reconstruct the microbial community encompassing the FC cecum. A total of 26 samples were analysed, involving samples that exhibited a typical FC-type microbiota based on 16S microbial profiling results and PCoA analysis (Fig. 1D). Selective and un-selective growth conditions based on the use of different modified media as well as various growth parameters such as 0_2 availability, antibiotics, micro- and macro- nutrient availability, fresh-sample and frozen sample inoculation, were applied to assess the microbial composition of these 26 cecal samples. About 1000 colonies were picked from plates prepared for all tested cultivation conditions. Furthermore, each of these colonies was subjected to additional purification processes and taxonomic identification based on 16S ribosomal RNA gene sequencing. Among the different cultivation conditions tested, the most effective, which resulted in the isolation of about 40% of the total number of identified bacterial species-based on the microbiota analysis, was cultivation on PYG in an anaerobic atmosphere followed by incubation for five days.

These analyses allowed the reconstruction of part of the FC core microbiota (FC-microbiota) consisting of 417 strains representing 43 different species, representative of 15 genera and belonging to the phyla *Actinobacteria*, *Bacteroidetes*, *Firmicutes* and *Proteobacteria* (Fig. 4). The FC-microbiota isolates were dominated by *Firmicutes*, reflecting the abundant presence of this phylum in the poultry intestine. In detail, *Lactobacillus* spp., *Enterococcus* spp. and *Escherichia* spp. were the most frequently isolated bacterial genera of FC-microbiota, representing 53%, 10% and 2.5% of the isolates respectively. However,

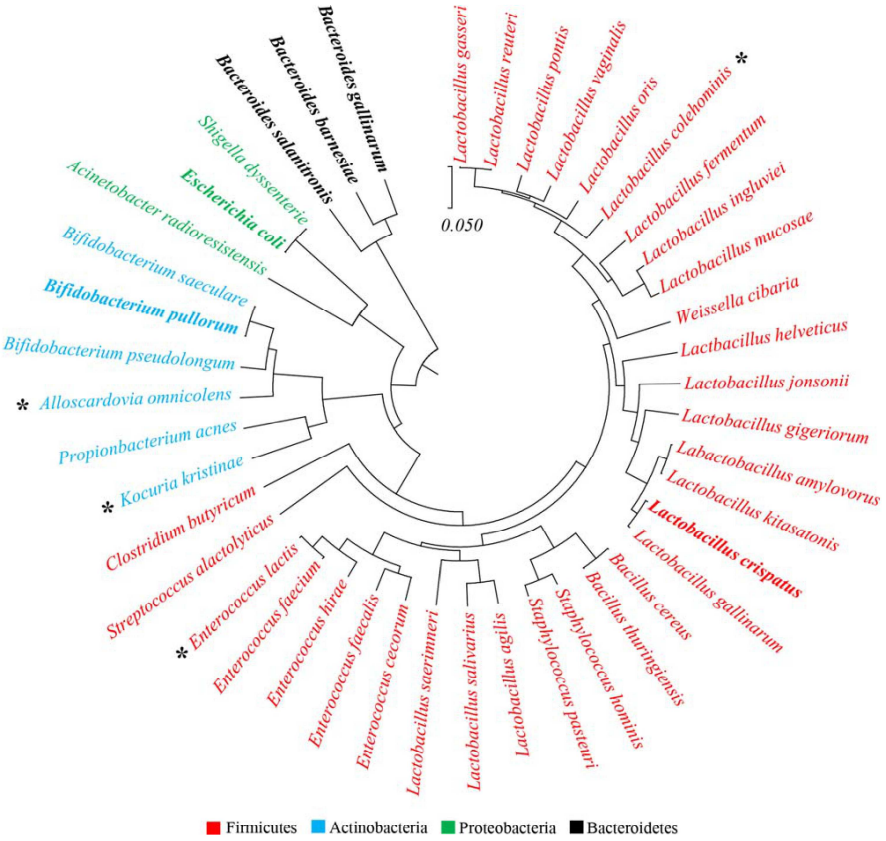


Fig. 4. Diversity of the FC poultry microbiota strain collection.
The picture shows the UPGMA phylogenetic tree based on the analysis of the sequenced 16S rRNA gene illustrating the classification of the isolated FC poultry microbiota members at species level. Only unique sequences were used. The tree is colour coded according to the phyla. Non-typical poultry isolates were marked with an asterisk. Bacterial taxa identified by culturomic and shotgun metagenomic approaches were reported in bold. The tree was generated using MEGA v. 7.0.2.

only a small number of isolates were classified as belonging to the *Bacteroidetes* phylum (Fig. 4). Arguably, conservation, manipulation or cultivation procedures of samples have impacted on the persistence of members of this bacterial taxon, apparently being particularly sensitive to the applied laboratory growth conditions, thus reducing the variety of isolated bacterial species.

The overall biodiversity of the FC-microbiota based on these culturomic analyses allowed us to reconstruct the phylogenetic tree shown in Fig. 4.

Interestingly, of the 420 strains forming the FC-microbiota, four taxa, that is, *Alloscardovia omnicolens*, *Enterococcus lactis*, *Kokuria kristinae* and *Lactobacillus colehominis*, representing approximately 4% of the number of taxa of FC-microbiota, have never been isolated from poultry previously, while the other strains are typical inhabitants of the chicken intestine (Collado and Sanz, 2007; Stanley *et al.*, 2014).

To assess the relevance of the culturomic approach reported here, a comparison between the cultivation-based method and high-throughput shotgun sequencing data was performed. Only six of the 43 species identified with culturomic-based methods were detected in shotgun results (about 15% of the total species detected by culturomics). Interestingly, these six bacterial species were spread across the four phyla identified using culture-based methods (Fig. 4).

Notably, comparison of 16S rRNA microbial profiling data with those achieved with culture-based approaches, highlighted that 60% of the bacterial genera detected with the culturomics approach were also identified with molecular methods. In contrast, just 3% of the bacterial genera identified by the 16S rRNA microbial profiling approach were isolated from our cultures of poultry cecal samples.

These findings highlight a substantial discordance between the sets of bacteria identified by these two approaches, at both genus and species level (Lagier *et al.*, 2016). Indeed, the ability of high throughput cultivation-based approaches to isolate bacteria from environmental samples that are not detected in genomic and metagenomic studies is far from effective. At the same time, large-scale molecular studies have a detection limit that prevents the detection of microorganisms that are present below a particular threshold (Lagier *et al.*, 2012).

Conclusions

Farm animals that undergo feralization follow a reverse domestication process, where they may gain certain traits that reflect their ancestors, while maintaining others that had been selected by humans such as brain dimension and organization, plumage and animal size (Gering *et al.*, 2015; Callaway, 2016; Johnsson *et al.*, 2016). Such animals also represent an intriguing example of microbiota evolution subsequent to human influences, which include changes in environment, diet and host genetics. To investigate how feralization modulates the microbiome, *in silico* and *in vitro* approaches were used to decipher the cecal microbiota of FC animals originated from Bermuda (and compare them to domesticated BC).

Notably, despite the fact that these feral animals live on an island, are not exposed to human husbandry, and follow a 'natural' diet, the Bermuda Poultry, that is, Feral, Free-range broilers and layers, microbiota displayed a composition that still resembles that of FRC cecal samples collected in very different and distant geographical regions (Mancabelli *et al.*, 2016). Interestingly, the reconstructed microbiota of FRC and Bermuda animals were shown to be characterized by a higher abundance of strains belonging to the *Bacteroidetes* phylum, being clearly different from the cecal microbiota of domesticated animals (i.e., BC). Furthermore, analysing the cecal microbiota of FC, BC and FRC, 18 taxa belonging to the *Firmicutes* phylum were present in all samples, thus forming a proposed core microbiota. Although the low number of analysed FC cecal samples may represent a limitation of this study, the interesting results obtained prompted us to sequence total DNA of Bermuda Feral cecal samples through a shotgun metagenomics approach. The functional microbiota characterization highlighted a selection in FC microbiota of those bacterial genera that are less efficient for energy recovery (or host weight gain), when compared to BC, yet carry a lower number of genes related to antibiotic resistance.

In addition, we performed a first attempt to reconstruct the cecal microbiota of FC. Culturomics approaches were applied to *in vitro* reconstruct the microbiota of FC animals, providing comprehensive culture conditions simulating or mimicking the environmental conditions present in the cecum of chickens. This approach enabled a large-scale cultivation of bacteria and the generation of a unique strain collection of about 420 isolates, including bacterial strains that are not typically retrieved from the chicken gut, and representing the four major bacterial phyla present in the cecal microbiota of feral poultry. Many efforts will still be needed to obtain a complete bacterial collection representing the complexity of the poultry microbiota, but these 15 samples are a starting point to understand the evolution of this bacterial community. Additional feral samples collected from other geographical locations as well as further growth media/cultivation conditions may result in a more complete culture collection of the feral chicken cecal microbiome.

Experimental procedures

Ethical statement

Collection and export of samples derived from feral chickens was approved by the Bermuda Department of Environment

and Natural Resources in support of the Bermuda Biodiversity Project. The samples manipulation and DNA extraction protocols were approved by the 'Comitato di Etica Università degli Studi di Parma', Italy. All procedures were performed in strict compliance with national guidelines (Decreto legislativo 26/2014) on the protection of animals used for scientific purposes.

Animals, sampling and DNA extraction

A total of 15 chickens from St. George's island, Bermuda, were investigated (Table 1). Ten animals were reared under natural conditions, without human control or interference, and therefore categorized as Bermuda Feral. Bermuda supports a large, self-sustaining population of feral chickens that have been living in the wild since at least the mid-1980s. Based on preliminary genetic and morphological analyses, these birds appear to be an admixed flock originating from several breeds that are popular sources of meat and eggs in the Western hemisphere. Our ad hoc observations of gut content suggest a highly variable diet including locally-occurring invertebrates (e.g., snails and insects), local ornamental and/or natural vegetation (e.g., seeds and shoots) and garbage from Bermuda households and businesses. We also sampled five birds from a Bermuda farm that were categorized as FRC since some human impact (e.g., shelter, ad libitum formulated diet and vaccination) was present. Two of these (BER 100–101, Free-Range Bermuda Broilers) were broiler breed males raised in a common, large outdoor enclosure and fed an ad libitum diet of growth-formulated feed. Three female brown pullets (BER 102–104, Free-Range Bermuda Layers) were also included that were housed in coops with an outdoor run, immediately adjacent to the broilers. All of the individuals from the Bermuda farm originated from a hatchery on the mainland US, where they were vaccinated for Marek's Disease Virus.

Cecal samples were obtained, kept under anaerobic conditions, transferred to the laboratory and maintained at −80°C until processing. We selected the ceca as organs of particular interest as they harbour the highest microbial cell densities (up to 10^{11} cells g^{-1}), have the longest residence time (12–20 h) of digesta in the gastrointestinal tract, and are important sites for carbohydrate fermentations, water regulation and recycling of urea (Oakley *et al.*, 2014; Sergeant *et al.*, 2014; Waite and Taylor, 2014). For DNA extraction, after removal of the digesta, 0.2 g of sample, composed of parts of both cecal pouches, were removed and briefly washed with Ringer's solution (Sigma, Italy) to remove unattached or loosely attached bacteria from the walls. Samples were subjected to DNA extraction using the QIAamp DNA Stool Mini kit following the manufacturer's instructions (Qiagen, USA).

16S rRNA gene amplification and MiSeq sequencing

Partial 16S rRNA gene sequences were amplified from extracted DNA using the primer pair Probio_Uni and/Probio_Rev, which target the V3 region of the 16S rRNA gene sequence (Milani *et al.*, 2013). Illumina adapter overhang nucleotide sequences were added to the partial 16S rRNA gene-specific amplicons, which were further processed using the 16S Metagenomic Sequencing Library Preparation Protocol (Part #15044223 Rev. B – Illumina; see also below) as previously reported (Mancabelli *et al.*, 2016). Purified amplicons were diluted to 4 nM and 5 μl aliquots of each diluted DNA amplicon were mixed to prepare the pooled final library. Sequencing was performed using an Illumina MiSeq sequencer with MiSeq Reagent Kit v3 chemicals.

16S rRNA-microbial profiling analysis

The fastq files were processed using QIIME (Caporaso *et al.*, 2010) as previously described (Milani *et al.*, 2013). Paired-end reads were merged and quality control retained sequences with a length between 140 and 400 bp, mean sequence quality score > 25 and with truncation of a sequence at the first base if a low quality 10 bp rolling window was found. Sequences with mismatched forward and/or reverse primers were omitted.

Shotgun metagenomics

DNA was fragmented to 550–650 bp using a BioRuptor machine (Diagenode, Belgium). Samples were prepared following the TruSeq Nano DNA Sample Preparation Guide (Part#15041110Rev.D). Sequencing was performed using an Illumina NextSeq 500 sequencer with NextSeq Mid Output v2 Kit chemicals.

Analysis of metagenomic data sets

The generated fastq files were filtered for reads with a quality score of < 25, for sequences of chicken genomic DNA, as well as for reads < 80 bp. Bases were also removed from the end of the reads unless the average quality score in a window of 5 bp was > 25. Only paired data were further analysed. The revised gene/protein set was searched using evolutionary genealogy of genes: Non-supervised Orthologous Groups (eggNOG; http://eggnog.embl.de/version_4.0.beta/) databases. Interrogation of sequence reads for significant identity to known antibiotic resistance genes (ARGs) was performed using a custom script based on RapSearch2 software (Zhao *et al.*, 2012), htseq-count (Anders *et al.*, 2015) and the database CARD (McArthur *et al.*, 2013), which encompasses amino acid sequences of enzymes involved in antibiotic resistance. Reconstruction of glycosyl hydrolase profiles and bacterial metabolic pathways, and evaluation of their abundance in the shotgun metagenomic data sets was performed using custom scripts based on RapSearch2 software (Zhao *et al.*, 2012), htseq-count (Anders *et al.*, 2015) and the CAZy database or the MetaCyc database (Caspi *et al.*, 2012) respectively.

In vitro *reconstruction of poultry microbiota*

For microbiota cultivation, different media were assessed, including Proteose Yeast Glucose (PYG) agar modified media (Leibniz Institut DSMZ-Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH) with modifications (0.5% resistant starch, 0.05% pectin from citrus peel and 0.25% inulin), YCFA agar (Duncan *et al.*, 2002) supplemented with 0.2% of glucose, maltose and cellobiose, brain heart infusion (BHI) agar modified (10 mg L^{-1} hemin, and 1 mg L^{-1} vitamin K, 0.05% L-

cysteine hydrochloride), de Man, Rogosa and Sharpe (MRS) agar supplemented with 0.05% L-cysteine hydrochloride and Chicken cecal medium (CCM), based on VL broth prepared as reported by Nisbet and colleagues (1993) and added with a chicken cecal content solution (autoclaved at 121°C for 15 min, 5% w/v) at a final concentration of 1% w/v. Finally, PYG medium was also supplemented with streptomycin (16 μg ml^{-1}) or vancomycin (2 μg ml^{-1}) after autoclaving.

All reagents were placed in the anaerobic cabinet (Ruskin, in which the atmosphere consisted of 17% CO_2, 80% N_2 and 2.99% H_2) for 24 h prior to use in order to create anaerobic conditions.

Cecal samples were obtained from animals categorized as FC-group members after 16S rRNA microbial profiling. All samples were processed immediately after their arrival at the laboratory.

About 0.2 g of cecum samples were washed with phosphate buffer solution (pH = 7.0), 10-fold diluted and then inoculated in all media previously described. Plates were incubated at 37°C for 5 days in anaerobic conditions, in 5% CO_2 and in aerobic conditions. Colonies were randomly picked, pure-cultured and DNA was extracted from each isolate through rapid mechanic cell lysis as described previously (Turroni *et al.*, 2009). Isolates from each sample were stored at −80°C in the presence of glycerol (30%, v/v).

Taxonomic identification of FC isolates

&&Identification of each isolate was performed by PCR amplification of a portion of the 16S rRNA gene using primers P0 (5′-GAAGAGTTTGATCCTGGCTCAG-3′) and P6 (5′-CTACGGC TACCTTGTTACGA-3′). Each 25 μl PCR reaction contained approximately 30 ng of genomic DNA, Platinum PCR SuperMix 1X (Invitrogen, USA) and 100 pM of each oligo. PCR reactions were performed on a Verity Thermocycler (Applied Biosystems, USA). Electrophoretic profiles were visualized by SYBR Safe DNA gel stain (Invitrogen). PCR product purification was performed using the NucleoSpin Gel and PCR Clean-up kit (Macherey-Nagel) following manufacturer instructions.

Each 16S rRNA gene thus generated from individual colonies originating from fecal samples was sequenced and it was then subjected to a BLAST search against the GenBank database.

Statistical analyses

ANOVA and PERMANOVA analyses were performed using Tuckey HSD post hoc test. All statistical analyses were performed with SPSS software (www.ibm.com/software/it/analytics/spss/).

Data deposition

16S rRNA-based microbial profiling data sets obtained in this study were deposited in SRA under accession numbers SRP114489. Shotgun metagenomic data sets are accessible through SRA study accession number SRP114492.

Acknowledgements

This is Contribution #260, Bermuda Biodiversity Project (BBP), Bermuda Aquarium, Natural History Museum and Zoo, Department of Environment and Natural Resources. This work was funded by the EU Joint Programming Initiative – A Healthy Diet for a Healthy Life (JPI HDHL, http://www.healthydietforhealthylife.eu/) to D.v.S. (in conjunction with Science Foundation Ireland (SFI), Grant number 15/JP-HDHL/3280) and to M.V. (in conjunction with MIUR, Italy). We thank GenProbio srl for financial support of the Laboratory of Probiogenomics. L.M. is supported by Fondazione Cariparma, Parma, Italy. D.v.S. is a member of The APC Microbiome Institute funded by Science Foundation Ireland (SFI), through the Irish Government's National Development Plan (Grant number SFI/12/RC/2273). EG was funded by the National Science Foundation under Cooperative Agreement No. DBI-0939454. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. The authors declare that they have no competing interests. Part of this research is conducted using the High Performance Computing (HPC) facility of the University of Parma. We thank Mark Outerbridge and Edward Wright for help with fieldwork and logistics.

References

Anders, S., Pyl, P.T., and Huber, W. (2015) HTSeq–a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31:** 166–169.

Astudillo-Garcia, C., Bell, J.J., Webster, N.S., Glasl, B., Jompa, J., Montoya, J.M., and Taylor, M.W. (2017) Evaluating the core microbiota in complex communities: a systematic investigation. *Environ Microbiol* **19:** 1450–1462.

Browne, H.P., Forster, S.C., Anonye, B.O., Kumar, N., Neville, B.A., Stares, M.D., *et al.* (2016) Culturing of 'unculturable' human microbiota reveals novel taxa and extensive sporulation. *Nature* **533:** 543–546.

Callaway, E. (2016) When chickens go wild. *Nature* **529:** 270–273.

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.

Caspi, R., Altman, T., Dreher, K., Fulcher, C.A., Subhraveti, P., Keseler, I.M., *et al.* (2012) The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. *Nucleic Acids Res* **40:** D742–D753.

Clemente, J.C., Ursell, L.K., Parfrey, L.W., and Knight, R. (2012) The impact of the gut microbiota on human health: an integrative view. *Cell* **148:** 1258–1270.

Collado, M.C., and Sanz, Y. (2007) Characterization of the gastrointestinal mucosa-associated microbiota of pigs and chickens using culture-based and molecular methodologies. *J Food Prot* **70:** 2799–2804.

den Besten, G., van Eunen, K., Groen, A.K., Venema, K., Reijngoud, D.J., and Bakker, B.M. (2013) The role of short-chain fatty acids in the interplay between diet, gut microbiota, and host energy metabolism. *J Lipid Res* **54:** 2325–2340.

Duncan, S.H., Hold, G.L., Harmsen, H.J., Stewart, C.S., and Flint, H.J. (2002) Growth requirements and fermentation products of Fusobacterium prausnitzii, and a proposal to reclassify it as Faecalibacterium prausnitzii gen. nov., comb. nov. *Int J Syst Evol Microbiol* **52:** 2141–2146.

Fraher, M.H., O'toole, P.W., and Quigley, E.M. (2012) Techniques used to characterize the gut microbiota: a guide for the clinician. *Nat Rev Gastroenterol Hepatol* **9:** 312–322.

Gering, E., Johnsson, M., Willis, P., Getty, T., and Wright, D. (2015) Mixed ancestry and admixture in Kauai's feral chickens: invasion of domestic genes into ancient Red Junglefowl reservoirs. *Mol Ecol* **24:** 2112–2124.

Hiergeist, A., Glasner, J., Reischl, U., and Gessner, A. (2015) Analyses of intestinal microbiota: culture versus sequencing. *ILAR J* **56:** 228–240.

Hu, Y., Yang, X., Qin, J., Lu, N., Cheng, G., Wu, N., *et al.* (2013) Metagenome-wide analysis of antibiotic resistance genes in a large cohort of human gut microbiota. *Nat Commun* **4:** 2151.

Huerta-Cepas, J., Szklarczyk, D., Forslund, K., Cook, H., Heller, D., Walter, M.C., *et al.* (2016) eggNOG 4.5: a hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences. *Nucleic Acids Res* **44:** D286–D293.

Human Microbiome Project Consortium. (2012) Structure, function and diversity of the healthy human microbiome. *Nature* **486:** 207–214.

Johnsson, M., Gering, E., Willis, P., Lopez, S., Van Dorp, L., Hellenthal, G., *et al.* (2016) Feralisation targets different genomic loci to domestication in the chicken. *Nat Commun* **7:** 12950.

Lagier, J.C., Million, M., Hugon, P., Armougom, F., and Raoult, D. (2012) Human gut microbiota: repertoire and variations. *Front Cell Infect Microbiol* **2:** 136.

Lagier, J.C., Hugon, P., Khelaifia, S., Fournier, P.E., La Scola, B., and Raoult, D. (2015) The rebirth of culture in microbiology through the example of culturomics to study human gut microbiota. *Clin Microbiol Rev* **28:** 237–264.

Lagier, J.C., Khelaifia, S., Alou, M.T., Ndongo, S., Dione, N., Hugon, P., *et al.* (2016) Culture of previously uncultured members of the human gut microbiota by culturomics. *Nat Microbiol* **1:** 16203.

Ley, R.E., Hamady, M., Lozupone, C., Turnbaugh, P.J., Ramey, R.R., Bircher, J.S., *et al.* (2008) Evolution of mammals and their gut microbes. *Science* **320:** 1647–1651.

Lozupone, C.A., Stombaugh, J.I., Gordon, J.I., Jansson, J.K., and Knight, R. (2012) Diversity, stability and resilience of the human gut microbiota. *Nature* **489:** 220–230.

Mancabelli, L., Ferrario, C., Milani, C., Mangifesta, M., Turroni, F., Duranti, S., *et al.* (2016) Insights into the biodiversity of the gut microbiota of broiler chickens. *Environ Microbiol* **18:** 4727–4738.

Mao, S., Zhang, M., Liu, J., and Zhu, W. (2015) Characterising the bacterial microbiota across the gastrointestinal tracts of dairy cattle: membership and potential function. *Sci Rep* **5:** 16116.

Marchesi, J.R., Adams, D.H., Fava, F., Hermes, G.D., Hirschfield, G.M., Hold, G., *et al.* (2016) The gut microbiota and host health: a new clinical frontier. *Gut* **65:** 330–339.

McArthur, A.G., Waglechner, N., Nizam, F., Yan, A., Azad, M.A., Baylay, A.J., *et al.* (2013) The comprehensive antibiotic resistance database. *Antimicrob Agents Chemother* **57:** 3348–3357.

McCormack, U.M., Curiao, T., Buzoianu, S.G., Prieto, M.L., Ryan, T., Varley, P., *et al.* (2017) Exploring a possible link between the intestinal microbiota and feed efficiency in pigs. *Appl Environ Microbiol* **83:**

Milani, C., Hevia, A., Foroni, E., Duranti, S., Turroni, F., Lugli, G.A., *et al.* (2013) Assessing the fecal microbiota: an optimized ion torrent 16S rRNA gene-based analysis protocol. *PLoS One* **8:** e68739.

Mirzaei, M.K., and Maurice, C.F. (2017) Menage a trois in the human gut: interactions between host, bacteria and phages. *Nat Rev Microbiol* **15:** 397–408.

Mohd Shaufi, M.A., Sieo, C.C., Chong, C.W., Gan, H.M., and Ho, Y.W. (2015) Deciphering chicken gut microbial dynamics based on high-throughput 16S rRNA metagenomics analyses. *Gut Pathog* **7:** 4.

Nisbet, D.J., Corrier, D.E., and DeLoach, J.R. (1993) Effect of mixed cecal microflora maintained in continuous culture and of dietary lactose on Salmonella typhimurium colonization in broiler chicks. *Avian Dis* **37:** 528–535.

Oakley, B.B., Lillehoj, H.S., Kogut, M.H., Kim, W.K., Maurer, J.J., Pedroso, A., *et al.* (2014) The chicken gastrointestinal microbiome. *FEMS Microbiol Lett* **360:** 100–112.

Osterblad, M., Norrdahl, K., Korpimaki, E., and Huovinen, P. (2001) Antibiotic resistance. How wild are wild mammals? *Nature* **409:** 37–38.

Roto, S.M., Rubinelli, P.M., and Ricke, S.C. (2015) An introduction to the avian gut microbiota and the effects of yeast-based prebiotic-type compounds as potential feed additives. *Front Vet Sci* **2:** 28.

Salonen, A., Salojarvi, J., Lahti, L., and de Vos, W.M. (2012) The adult intestinal core microbiota is determined by analysis depth and health status. *Clin Microbiol Infect* **18(Suppl 4):** 16–20.

Salyers, A.A., Gupta, A., and Wang, Y. (2004) Human intestinal bacteria as reservoirs for antibiotic resistance genes. *Trends Microbiol* **12:** 412–416.

Sergeant, M.J., Constantinidou, C., Cogan, T.A., Bedford, M.R., Penn, C.W., and Pallen, M.J. (2014) Extensive microbial and functional diversity within the chicken cecal microbiome. *PLoS One* **9:** e91941.

Stanley, D., Geier, M.S., Denman, S.E., Haring, V.R., Crowley, T.M., Hughes, R.J., and Moore, R.J. (2013) Identification of chicken intestinal microbiota correlated with the efficiency of energy extraction from feed. *Vet Microbiol* **164:** 85–92.

Stanley, D., Hughes, R.J., and Moore, R.J. (2014) Microbiota of the chicken gastrointestinal tract: influence on health, productivity and disease. *Appl Microbiol Biotechnol* **98:** 4301–4310.

Thaller, M.C., Migliore, L., Marquez, C., Tapia, W., Cedeno, V., Rossolini, G.M., and Gentile, G. (2010) Tracking acquired antibiotic resistance in commensal bacteria of Galapagos land iguanas: no man, no resistance. *PLoS One* **5:** e8989.

Turroni, F., Foroni, E., Pizzetti, P., Giubellini, V., Ribbera, A., Merusi, P., *et al.* (2009) Exploring the diversity of the bifidobacterial population in the human intestinal tract. *Appl Environ Microbiol* **75:** 1534–1545.

von Wintersdorff, C.J., Penders, J., van Niekerk, J.M., Mills, N.D., Majumder, S., van Alphen, L.B., *et al.* (2016) Dissemination of antimicrobial resistance in microbial ecosystems through horizontal gene transfer. *Front Microbiol* **7:** 173.

Waite, D.W., and Taylor, M.W. (2014) Characterizing the avian gut microbiota: membership, driving influences, and potential function. *Front Microbiol* **5:** 223.

Weinstock, G.M. (2012) Genomic approaches to studying the human microbiota. *Nature* **489:** 250–256.

Wright, G.D. (2007) The antibiotic resistome: the nexus of chemical and genetic diversity. *Nat Rev Microbiol* **5:** 175–186.

Yassour, M., Vatanen, T., Siljander, H., Hamalainen, A.M., Harkonen, T., Ryhanen, S.J., *et al.* (2016) Natural history of the infant gut microbiome and impact of antibiotic treatment on bacterial strain diversity and stability. *Sci Transl Med* **8:** 343ra381.

Zhao, Y., Tang, H., and Ye, Y. (2012) RAPSearch2: a fast and memory-efficient protein similarity search tool for next-generation sequencing data. *Bioinformatics* **28:** 125–126.

Supporting information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Fig. S1. 16S rRNA-microbial profiling of the chicken samples at phylum level, encompassing BC, FRC and FC animals. Sample categorization in FRC, BC or FC was performed subsequent to obtaining the results from the PCoAs as reported in Figure 1c and d. U. m.: unclassified member.

Fig. S2. 16S rRNA-microbial profiling of the chicken samples at family level, encompassing BC, FRC and FC animals. Sample categorization in FRC, BC or FC was performed subsequent to obtaining the results from the PCoAs as reported in Figure 1c and d. U. m.: unclassified member.

Fig. S3. 16S rRNA-microbial profiling of the chicken samples at genus level, encompassing BC, FRC and FC animals. Sample categorization in FRC, BC or FC was performed subsequent to obtaining the results from the PCoAs as reported in Figure 1c and d. Only taxa with a relative abundance of $>0.5\%$ are shown. U. m.: unclassified member.

Fig. S4. Taxonomic data at phylum level obtained from shotgun metagenomics of FC and BC cecal samples.

Table S1. 16S rRNA microbial profiling data.

Table S2. Bacterial taxonomy from shotgun metagenomics sequencing of FC and BC.

Table S3. Functional annotation of FC metagenomic data sets according to COG categories. Each COG family is identified by a one-letter abbreviation

Table S4. Carbohydrate degradation related enzymes of FC microbiomes.

Table S5. Presence of pathways involved in formate, acetate, propionate and butyrate production in FC microbiomes.

Table S6. Presence of predicted enzymes involved in antibiotic resistance in FC shotgun metagenomics data set.

Chapter 5

Meta-analysis of the human gut microbiome from urbanized and pre-agricultural populations

Mancabelli L, Milani C, Lugli GA, Turroni F, Ferrario C, van Sinderen D, Ventura M.

This chapter was published in:

Environmental Microbiology. 2017 Apr;19(4):1379-1390. Epub 2017 Feb 28. Review.

environmental microbiology

Environmental Microbiology (2017) **19**(4), 1379–1390 doi:10.1111/1462-2920.13692

Minireview

Meta-analysis of the human gut microbiome from urbanized and pre-agricultural populations

Leonardo Mancabelli,[1†] Christian Milani,[1†] Gabriele Andrea Lugli,[1] Francesca Turroni,[1] Chiara Ferrario,[1] Douwe van Sinderen,[2] Marco Ventura[1*]

[1]*Laboratory of Probiogenomics, Department of Chemistry, Life Sciences and Environmental Sustainability, University of Parma, Parma, Italy.*

[2]*APC Microbiome Institute and School of Microbiology, Bioscience Institute, National University of Ireland, Cork, Ireland.*

Summary

Metagenomic studies of the human gut microbiome have only recently begun to explore the differences in taxonomic composition between subjects from diverse geographical origins. Here, we compared taxonomy, resistome and functional metabolic properties of publicly available shotgun datasets of human fecal samples collected from different geographical regions (Europe, North America, Asia and Oceania). Such datasets encompassed gut microbiota information corresponding to 13 developed/industrialized societies, as well as two traditional hunter-gatherer, pre-agricultural communities (Tanzanian and Peruvian individuals). Assessment of the retrieved taxonomic profiles allowed the most updated reconstruction of the global core-microbiome as based on currently available data, as well as the identification and targeted genome reconstruction of bacterial taxa that appear to have been lost and/or acquired during urbanization/industrialization. Functional characterization of these metagenomic datasets indicates that the urbanization/industrialization process which occurred in recent human history has shaped the gut microbiota through the acquisition and/or loss of specific gut microbes, thereby potentially impacting on the overall functionality of the gut microbiome.

Received 25 October, 2016; revised 3 February, 2017; accepted 6 February, 2017. *For correspondence. E-mail: marco.ventura@unipr.it; Tel. +39-521-905666; Fax +39-521-905604. †These authors contributed equally.

Introduction

The composite activities of the human gut microbiome impact on various functions of its host, including gut physiology, intestinal metabolism, and immune system modulation (Round and Mazmanian, 2009). During the life span of its host, the gut microbiota composition is influenced by factors such as diet, lifestyle and environment (Conlon and Bird, 2015). Recently, some major research efforts, such as the European Metagenomics of the Human Intestinal Tract (MetaHIT) (http://www.metahit.eu) project and the American Human Microbiome Project (HMP) (http://hmpdacc.org), have dissected the gut microbiome composition and functionality across different human populations. In order to better understand the role of the microbiota and its co-evolution with human host, it's pivotal to compare the microbiome of urbanized/industrialized populations with pre-agricultural/isolated populations. Currently, only a small number of studies have compared the microbiomes of pre-agricultural/isolated communities and urbanized/industrialized populations in order to detect possible differences in composition and their potential correlation with disease (risk) and/or metabolic disorders (De Filippo *et al.*, 2010; Yatsunenko *et al.*, 2012; Schnorr *et al.*, 2014; Clemente *et al.*, 2015; Martinez *et al.*, 2015; Obregon-Tito *et al.*, 2015; Rampelli *et al.*, 2015).

Moreover, many investigations of the gut microbiota were simply interested in microbial cataloguing by means of 16S rRNA gene-based amplicon sequencing (De Filippo *et al.*, 2010; Yatsunenko *et al.*, 2012; Schnorr *et al.*, 2014; Clemente *et al.*, 2015; Dehingia *et al.*, 2015; Martinez *et al.*, 2015). Nonetheless, an increasing number of investigative efforts are based on shotgun metagenomics sequencing, aimed at functionally characterizing gut microbiomes (Qin *et al.*, 2010; Qin *et al.*, 2012; Karlsson *et al.*, 2013; Li *et al.*, 2014; Lim *et al.*, 2014; Zeller *et al.*, 2014; Feng *et al.*, 2015; Obregon-Tito *et al.*, 2015; Rampelli *et al.*, 2015; Voigt

et al., 2015). The majority of gut microbiota analyses have focused on specific populations, yet do not provide any comparative analysis of gut microbiomes from different geographical regions or habitation conditions (Karlsson *et al.*, 2013; Lim *et al.*, 2014; Feng *et al.*, 2015; Voigt *et al.*, 2015; Raymond *et al.*, 2016).

So far, only a small number of studies have investigated how the urbanization/industrialization process may have affected gut microbiota composition (De Filippo *et al.*, 2010; Yatsunenko *et al.*, 2012; Schnorr *et al.*, 2014; Clemente *et al.*, 2015; Dehingia *et al.*, 2015; Martinez *et al.*, 2015; Obregon-Tito *et al.*, 2015; Rampelli *et al.*, 2015). However, most of these investigations were interested in the identification of specific microbial taxa that correlate with the adaptation of human to different lifestyles without supporting their findings with shotgun metagenomics analyses (De Filippo *et al.*, 2010; Yatsunenko *et al.*, 2012; Schnorr *et al.*, 2014; Clemente *et al.*, 2015; Dehingia *et al.*, 2015; Martinez *et al.*, 2015). These studies reported a range of differences in the gut microbiota composition between industrialized and pre-agricultural societies, reflecting the dietary and environmental factors typical of their lifestyle. Particularly, pre-agricultural communities displayed high abundance of members of the *Prevotella* genus, known to harbour genetic features for the breakdown of cellulose and xylan. While the taxa *Treponema* and *Brachyspira* were undetected in industrialized populations (Schnorr *et al.*, 2014; Obregon-Tito *et al.*, 2015; Rampelli *et al.*, 2015). Members of the genus *Treponema* were also found in non-human primates and all traditional populations studied to date, suggesting that these gut commensals have been lost from the gut microbiota of human beings associated with urban-industrialized societies (Obregon-Tito *et al.*, 2015).

So far, only two publications have reported on the use of shotgun sequencing to perform an in depth functional analysis of the gut microbiome of Hadza and Matses populations, which represent pre-agricultural societies (Obregon-Tito *et al.*, 2015; Rampelli *et al.*, 2015). These studies reported that Hazda and Matses microbiomes possess a lower abundance of antibiotic resistance genes and an extended metabolic potential toward utilization of carbohydrates when compared with human gut microbiomes from industrialized areas (Obregon-Tito *et al.*, 2015; Rampelli *et al.*, 2015). Nevertheless, both studies compared the gut microbiomes of pre-agricultural populations to very small cohorts of fecal samples collected from individuals living in urbanized societies that can not be considered representative of the general western population (Obregon-Tito *et al.*, 2015; Rampelli *et al.*, 2015).

In this minireview, we evaluated the notion that urbanization/industrialization processes have subtantially influenced the composition and functionality of the human gut microbiome. This evaluation was performed from a taxonomic and functional perspective by means of a meta-analysis of all publicly available human gut shotgun metagenomic datasets corresponding to urbanized and pre-agricultural societies. An overall summary of the observations discussed in this study, accompained by a comparison with previously published data, is reported in Supporting Information Table S1.

The urbanized/industrialized populations include individuals residing in high-income geographical regions that are densely populated, i.e. metropolitan areas, towns, but also individuals inhabiting rural areas that have access to medical care and obtain high hygiene standards and follow a globalized westernized diet (Karlsson *et al.*, 2013; Li *et al.*, 2014; Lim *et al.*, 2014; Zeller *et al.*, 2014; Voigt *et al.*, 2015; Raymond *et al.*, 2016). In contrast, the pre-agricultural communities encompass individuals living in isolated areas, who have no or limited access to medical care, and whose diet is based on foods gathered and/or huntered from their immediate environment, which was further processed only by cooking (Schnorr *et al.*, 2014; Obregon-Tito *et al.*, 2015). The lifestyle (including diet) of these human communities resembles that of people from ancestral human populations (Clemente *et al.*, 2015). We would also have liked to evaluate the microbiome of ancestral human populations and microbiota-host co-evolution at various stages of human evolution, from the Neolithic to the modern urbanized/industrilized populations. Unfortunately, such meta-studies that cover human evolutionary history are currently not possible due to a lack of corresponding data sets.

These data allow the identification of specific compositional and functional differences between the gut microbiome of urbanized/-industrialized vs. pre-agricultural populations. Furthermore, this metagenomic information may allow genome reconstruction of bacterial taxa that seem to have been lost from or gained by individuals living in urban-industrialized countries.

The worldwide gut microbiome database

All datasets included in this meta-analysis were collected from published human gut microbiome studies. Literature searches allowed us to exclusively select shotgun metagenomic data generated by Illumina technology, i.e. the (currently) most preferred and reliable technology to perform shotgun metagenomics studies (Quail *et al.*, 2012). Datasets had to be obtained starting from DNA extracted from fecal samples of healthy and adult human individuals. In order to obtain comparable sequence information, it was necessary to exclude those metagenomic projects that did not have at least six datasets with an average quality value of >25 and an average

read length of > 95 bp (following quality filtering). These criteria resulted in the selection of 18 publicly available shotgun metagenomic sequencing projects from 14 different geographical regions covering Africa, Asia, Europe, North/South America and Oceania (Supporting Information Table S2).

Metadata of the sequencing projects was employed to select only datasets of healthy individuals whose age ranged between 21 and 65, not undergoing any antibiotic or probiotic treatment, and not suffering from gut-related diseases/disorders. Unfortunately, the metadata of these collected datasets frequently do not provide information related to eating habits, diet and associated nutritional proprieties. This represents a crucial limitation which prevents us from assessing the role of diet in shaping the gut microbiota composition. In particular, human gut microbiomes were obtained for individuals from Austria, Denmark, France, Germany, Spain, Sweden, China, South Korea, Canada (Qin *et al.*, 2010; 2012; Karlsson *et al.*, 2013; Li *et al.*, 2014; Lim *et al.*, 2014; Zeller *et al.*, 2014; Feng *et al.*, 2015; Voigt *et al.*, 2015; Nishijima *et al.*, 2016; Raymond *et al.*, 2016), Italy (SRP079680), Australia (PRJEB6092) and United States (HMP DACC, http://www.hmpdacc.org). Furthermore, the gut microbiomes of just two pre-agricultural communities, i.e. Hadza (from Tanzania) (Schnorr *et al.*, 2014; Rampelli *et al.*, 2015) and Matses (from Peru) (Obregon-Tito *et al.*, 2015), were publicly avaiable and thus included in our analyses. While the avaibility of two datasets is not sufficient for a comprehensive and statistically significant representation of pre-agricultural populations in general, these data at least allow us to gain some initial insights into the effects of urbanization/industrialization on gut microbiome.

Samples were taxonomically profiled through MetaPhlAn2 software (Truong *et al.*, 2015) and the 10 samples with profiles closest to the average of each population were chosen as representatives for in-depth functional analyses. These datasets were included in the Worldwide Gut Microbiome Database (WGMD), and encompassed a total of 142 samples, facilitating both a taxonomic and a functional overview of the covered countries (Supporting Information Table S2).

Notably, the use of shotgun metagenomic data minimizes biases, in particular when compared with other taxonomic profiling methods that are based on marker genes, such as the widely used 16S rRNA profiling. In fact, for this latter approach there are serious methodological concerns (e.g. efficiency and universality of the PCR primers and DNA extraction protocols) (Milani *et al.*, 2013) that might prevent a faithful comparison between data sets derived from different studies. Moreover, the use of shotgun datasets allows us to perform functional investigation of the human gut microbiome, which is not possible with the use of 16S rRNA profiling data sets. Thus, despite limitations related to different procedures concerning stool collection/processing and DNA extraction protocols followed for the analysed samples (Qin *et al.*, 2010; 2012; Karlsson *et al.*, 2013; Li *et al.*, 2014; Lim *et al.*, 2014; Schnorr *et al.*, 2014; Zeller *et al.*, 2014; Feng *et al.*, 2015; Obregon-Tito *et al.*, 2015; Rampelli *et al.*, 2015; Voigt *et al.*, 2015; Nishijima *et al.*, 2016; Raymond *et al.*, 2016), the geographical coverage of the WGMD allows a global gut microbiota assessment related to both composition and functionality (Supporting Information Table S2).

Profiling of the global human gut microbiota

A 3-Dimensional Principal Coordinate Analysis (PCoA) representing the beta-diversity (Caporaso *et al.*, 2010) based on Bray-curtis dissimilarity index of the different gut microbiomes encompassing the WGMD, highlighted an intriguing profiling based on geographical regions (Fig. 1). In this context, the majority of the samples group together based on their geographic origin, whereas Australian, German and South Korean individuals elicit an uneven distribution (Fig. 1). Notably, the Hadza and Matses pre-agricultural individuals grouped as two separate clusters with respect to all other datasets, likely reflecting their distinct geographical origin, as well as their unique life style and diet (Supporting Information Table S3). Nevertheless, an extension of this meta-analysis with additional pre-agricultural samples is needed in order to statistically validate these results. Overall, the obtained results are statistically supported by a PERMANOVA p-value of < 0.001. Such findings may indicate that lifestyle and diet are important factors influencing the gut microbiota composition. In contrast, the urbanized populations show partial overlap, which may point towards the presence of a shared core microbiome (see below).

Inspection of predicted taxonomic profiles of the WGMD at phylum level show a preponderant presence of members of the *Bacteroidetes* phylum (Supporting Information Fig. S1), in particular of the *Bacteroides* genus in the urban-industrialized populations (average of 28.76%) (p-value < 0.05) (Supporting Information Fig. S2), whereas the *Prevotella* genus is highly abundant in the Hadza community (34.42%) (p-value < 0.05) (Supporting Information Fig. S2). Although some members of the *Bacteroides* genus are known to metabolize complex polysaccharides (Xu and Gordon, 2003), recent studies have shown that the presence of members of the *Bacteroides* genus positively correlates with a diet enriched in protein and animal fat. The presence of members of the *Prevotella* genus on the other hand is associated with regular consumption of a carbohydrate-based diet,

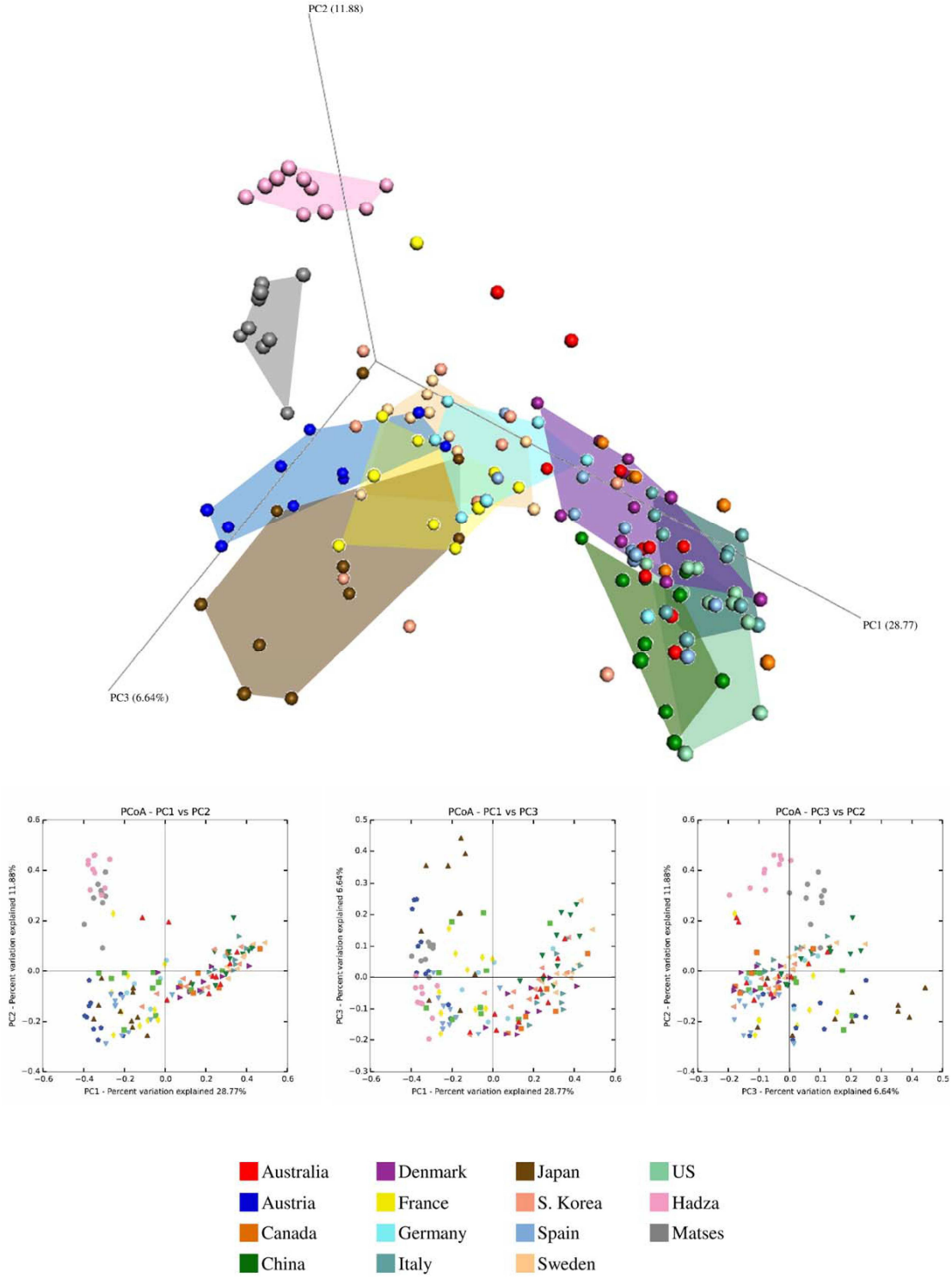


Fig. 1. Evaluation of the beta-diversity in the 142 analysed samples. The predicted PCoA is reported through two three-dimensional images as well as two-dimensional sections. The Panel depicts the beta-diversity of the samples subdivided according to their geographical origin. Colored areas highlight the main identified clusters. [Colour figure can be viewed at wileyonlinelibrary.com]

as commonly used in rural communities (Wu *et al.*, 2011).

Matses samples are characterized by low abundance of *Bacteroidetes* (3.91%) and high abundance of the phylum *Euryarchaeota* (15.03%), represented mainly by the *Methanobrevibacter* genus (14.90%) (Supporting Information Fig. S2). Members of this genus, such as *Methanobrevibacter smithii*, have been shown to increase the efficiency of energy extraction from dietary polysaccharides with consequent impact on host energy harvest (Samuel *et al.*, 2007). Therefore, *M. smithii* has been proposed as a target to reduce energy harvest in obese individuals (Samuel *et al.*, 2007).

Moreover, individuals from industrialized countries were shown to elicit a higher abundance of *Alistipes* genus (average of 5.12%) as compared with the two hunter-gatherer populations (average of 0.02%) (p-value < 0.01). This is in accordance with previous reports, suggesting a correlation between the *Bacteroides* enterotype and the presence of *Alistipes* (Wu *et al.*, 2011). Furthermore, the two hunter-gatherer communities show a higher abundance of the *Phascolarctobacterium* genus (average of 7.03%) compared with the abundance of this genus in individuals from industrialized nations (average of 0.35%) (p-value < 0.01), being notably absent from US, Swedish and Asian microbiome datasets. Members of the genus *Phascolarctobacterium* are known to produce high amounts of the short chain fatty acids (SCFA) acetate and propionate (Watanabe *et al.*, 2012). The majority of SCFA in the gut is derived from bacterial fermentation of complex carbohydrates present in this body compartment, such as dietary soluble fibres or resistant starch, which represent two of the main glycan components of the hunter-gatherer community diet (Watanabe *et al.*, 2012).

The 'pan-microbiome' of the human gut

Comparisons between different microbiomes are known to allow identification of bacterial taxa that are unique and thus characteristic of a specific metagenomic sample as well as those microbial groups that are commonly shared by all the microbiomes (Li *et al.*, 2014). Such analyses are thus important in the reconstruction of the so-called 'pan-microbiome' (Leung *et al.*, 2016). Interestingly, the pan-microbiome of WGMD reconstructed by means of MetaPhlAn2 software consists of 147 bacterial genera of which 48 were found to be present in $> 90\%$ of the samples in at least one population (Fig. 2).

The profiling data sets allow the identification of the human intestinal core microbiota, which represents bacteria that are present among all analysed human populations (Salonen *et al.*, 2012). While the number of core taxa will be reduced with the addition of future shotgun datasets covering new populations across the world, we believe that our iteration represents the most current and accurate reconstruction of the human core gut microbiota. These microorganisms are presumed to include evolutionary selected symbionts that strictly co-evolved with human beings and exert key roles in the biology of their host, such as added metabolic abilities, pathogen resistance and enhanced immune functionalities (Salonen *et al.*, 2012). Comparative analyses led to the identification of 22 genera present in all fecal samples of the 15 populations included in this study, thus representing the global human core microbiota (Fig. 2). Interestingly, in the urban-industrialized and pre-agricultural communities, this core microbiota encompass on average $82.13\% \pm 8.35\%$ and $46.95\% \pm 6.34\%$ (p-value < 0.01), respectively, of the total gut microbiota (Fig. 2). Furthermore, the analysis of the 'pan-microbiome' shows 15 accessory genera, which are not present in all populations. These taxa represent $11.63\% \pm 9.40\%$ and $39.65\% \pm 9.20\%$ of the microbiota in individuals living in urbanized-industrialized and pre-agricultural environments respectively (Fig. 2). Notably, these findings highlight that the gut microbiota of Hadza and Matses populations possess higher biodiversity, thus supporting the notion that urbanization/industrialization somehow caused a simplification of the microbial gut community (Schnorr *et al.*, 2014).

Furthermore, the pre-agricultural communities Hadza and Matses were found to be characterized by the presence of certain genera, that are absent in all other analysed populations, examples of which are *Treponema* and *Brachyspira* with average relative abundance of 4.95% and 0.04% respectively (Fig. 2). Notably, while *Treponema* is generally linked to infectious diseases (Giacani and Lukehart, 2014), a recent study reported that members of this genus may play a functional role in nutrient extraction from fibrous foods typically abundant in pre-agricultural diets (Schnorr *et al.*, 2014). Intriguingly, despite the limited availability of datasets from pre-agricultural populations, the present meta-analysis highlights that these taxa may represent microorganisms that were lost during urbanization/industrialization. Future integration of these datasets with additional shotgun metagenomics data from pre-agricultural populations will be necessary in order to confirm these results. To explore the functional roles exerted by these possibly lost gut microbiota members, targeted genome reconstruction and functional characterization of the main identified representative of this genus, i.e. *Treponema succinifaciens*, starting from shotgun metagenomics data collected for Hadza and Matses populations, allow us to obtain insights into the predicted metabolic properties of these bacteria(see below).

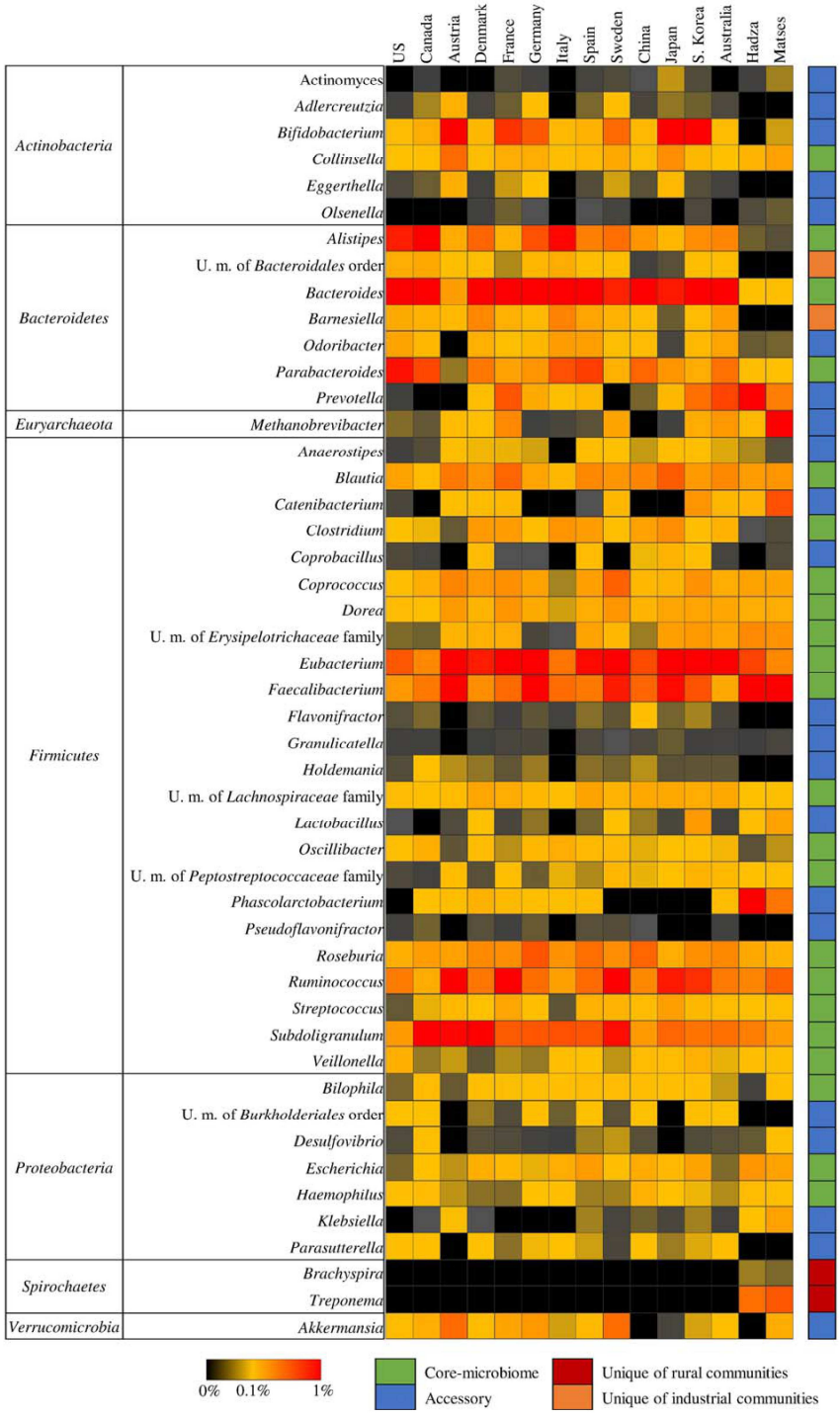


Fig. 2. The global human gut 'Pan-microbiome' at genus level. The heat map shows the relative abundance of bacterial taxa observed in 90% of the samples in at least one population. The right column indicates a classification of the taxa based on cell color: green cells represent taxa of the core-microbiome, blue cells indicate the accessory genera, and red and orange cells represent unique taxa of rural-primitive and urbanized-industrialized communities respectively. [Colour figure can be viewed at wileyonlinelibrary.com]

In contrast, the taxa Unclassified member of *Bacteroidales* order and *Barnesiella* (represented mainly by the species *Barnesiella intestinihominis*) are identified only in populations living in developed countries, suggesting that their presence was promoted by the urbanization/industrialization process.

Global microbiome functional characterization

The WGMD analysis also provided an updated snapshot of the global gut microbiome functionalities. *In silico* analyses of the predicted glycobiomes of shotgun metagenomic reads based on the Carbohydrate-Active enZYmes (CAZy) database, which includes the overall enzyme collection involved in glycan metabolism (Lombard *et al.*, 2014), allowed the detection of significant differences in relative abundance of Glycoside Hydrolase families (GH) between the communities included in the WGMD. Interestingly, Matses and Hadza communities show reduced abundance of GH-encoding reads (average of 1.18%) as compared with urban-industrialized populations (average of 2.92%) (Fig. 3a). Moreover, the microbiomes of hunter-gatherer communities possess a lower number of GH families (an average of 26) compared with those of the industrialized populations (an average of 46), based on GH families with an abundance of $>0.01\%$ of the total GH pool. A detailed analysis of these predicted glycobiomes, i.e., the enzymatic arsenal involved in the metabolism of carboydrates (Lombard *et al.*, 2014), revealed that industrialized communities encompass a higher abundance (an average of 1.35%) of GH families dedicated to the breakdown of multiple carbohydrate substrates (e.g. GH2, GH3 and GH43) compared with pre-agriculture microbiomes (with average of 0.50%) (p-value <0.05), which may be the consequence of the former group following a diet that is more diverse in glycan content. The described diet of pre-agricultural communities (Supporting Information Table S3) predominantly consists of a limited variety of tubers, vegetables and fruits gathered from the surrounding environment, while consumption of meat and eggs is infrequent (Schnorr *et al.*, 2014). Notably, the advent of agriculture and industrialization brought novel foods and advanced food processing technologies developed for food making as well as preservation and nutrition enhancement, which in turn led to new food and nutrient combinations that were absent in pre-agricultural human diets (Cordain *et al.*, 2005). Moreover, industrialization led to the development of global markets that further extended the variety of tubers, vegetables and fruits of the Western diet (Popkin *et al.*, 2012). In this context, our findings indicate that, even if the diet of pre-agricultural communities primarily consists of vegetables and fruits, it is limited in its glycan variety as compared with a typical Western diet (Cordain *et al.*, 2005), which may be responsible for a low GH diversity in the former case (Fig. 3a). This is in contrast with previous observations derived from limited comparisons between small cohorts of samples (Obregon-Tito *et al.*, 2015; Rampelli *et al.*, 2015) and highlights the critical role of a comprehensive and world-wide meta-analysis.

Pathway prediction of the WGMD gut microbiomes, as based on the MetaCyc database (Caspi *et al.*, 2012), show that, on average, $1.4\% \pm 0.47\%$ (Fig. 3b) of the microbiome data retrieved from pre-agricultural communities is involved in carbohydrate degradation (Supporting Information Table S4), constituting 30 different pathways with an abundance of $>0.01\%$. In contrast, gut microbiome metadata obtained from communities living in developed countries contained an average of $2.15\% \pm 0.65\%$ (Fig. 3b) reads encoding an average of 38 carbohydrate utilization pathways with an abundance of $>0.01\%$. Furthermore, Hadza and Matses populations contain on average a lower number of pathways involved in alginate (-98.68%), carrageenan (-100.00%) and gellan (-77.62%) degradation, as compared with urban populations. Alginate, carrageenan and gellan gum are three polysaccharides, which are derived from algae or bacteria, and which are widely used in the food industry as stabilizers, thickeners or emulsifying agents (Brownlee *et al.*, 2005; Burges Watson, 2008; Prajapati *et al.*, 2013).

Moreover, profiling of pathways involved in fatty acid and lipid metabolism indicates on average a lower level of triacylglycerol degradation in hunter-gatherer communities (-96.47%) as compared with industrialized populations, possibly reflecting the lower (animal-derived) fat intake of pre-agricultural populations (Supporting Information Table S3) (Cordain *et al.*, 2005).

These differences reinforce the notion that increased diet variety, as associated with an urbanized life style, has shaped the human microbiome towards an expansion of its carbohydrate degradative abilities.

In order to map and characterize the resistome of the microbial consortia residing in the gut of humans living in different geographic regions, the shotgun metagenomic datasets encompassing the WGMD were also screened for known bacterial Antibiotic Resistance Enzymes (AREs) (McArthur *et al.*, 2013) (Fig. 3c). *In silico* analyses show a high abundance of AREs in Japanese ($0.49\% \pm 0.10\%$), Korean ($0.30\% \pm 0.02\%$), Austrian ($0.32\% \pm 0.03\%$), Swedish ($0.31\% \pm 0.02\%$) and Italian individuals ($0.43\% \pm 0.05\%$), while the remaining eight industrialized countries display a somewhat lower average ARE abundance of $0.25\% \pm 0.03\%$ (p-value <0.05) (Fig. 3c). Interestingly, the data obtained from Tanzanian and Peruvian individuals indicate a significantly lower abundance of AREs (average of 0.19%)

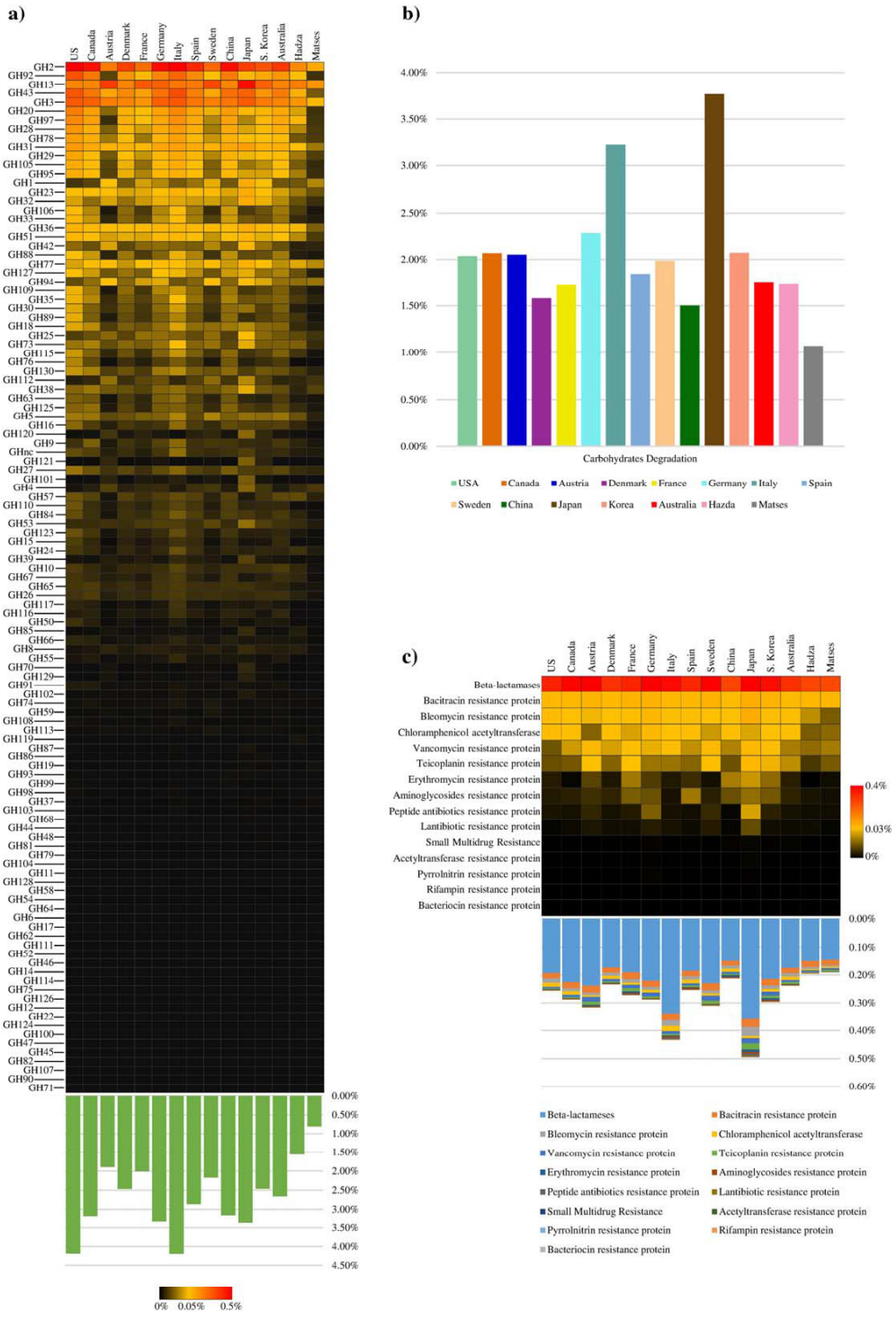


Fig. 3. Functional profile of the global human gut microbiome. Panel a shows a heat map reporting the abundance of glycosyl hydrolases in the various datasets, with a bar plot of their total abundance at the bottom. Panel b reports the % of the microbiome encoding pathways involved in carbohydrate degradation for each nation. Panel c exhibits AREs in the analysed populations, with a bar plot of the total abundance of these enzymes reported at the bottom. [Colour figure can be viewed at wileyonlinelibrary.com]

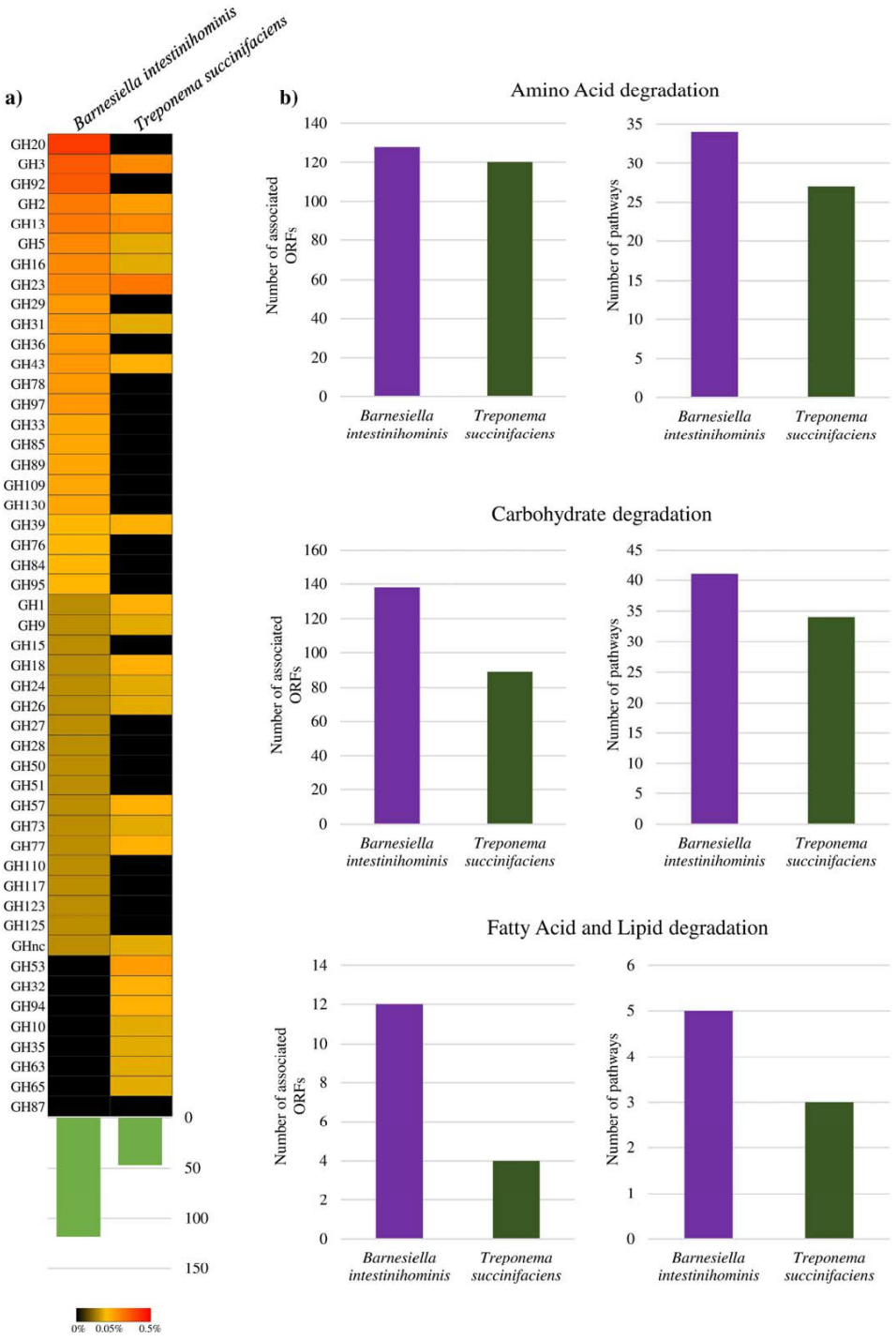


Fig. 4. Functional analysis of the reconstructed genome of *Barnesiella intestinihominis* and *Treponema succinifaciens*. Panel a indicates the abundance of glycoside hydrolases in *B. intestinihominis* and *T. succinifaciens*. The total number of GH-encoding genes is reported at the bottom. Panel b displays the number of genes and pathways involved in Amino Acid degradation, Carbohydrate degradation and Fatty Acid and Lipid degradation predicted in the analysed *B. intestinihominis* and *T. succinifaciens* genomes. [Colour figure can be viewed at wileyonlinelibrary.com]

compared with urban populations (p-value < 0.05). These results are in accordance with the very rare use of antibiotics by these pre-agricultural communities, consistent with a previous study (Clemente *et al.*, 2015).

Moreover, β-lactamase-encoding genes are the most abundant in all analysed samples, ranging from 0.15% ± 0.01% to 0.36% ± 0.07% in Matses and Japanese populations, respectively, possibly reflecting the fact that β-lactam-based antibiotics are the most widespread in both nature and common medical practice (Zeng and Lin, 2013).

Genome reconstruction of bacterial strains from shotgun metagenomic data

A valuable approach to assess functional roles exerted by bacterial taxa that were lost and/or acquired during urbanization/industrialization, is the targeted genome reconstruction [through the MEGAnnotator pipeline (Lugli *et al.*, 2016)] of unique gut commensals identified in the various gut microbiomes of WGMD. Such analyses targeted the reconstruction of *Treponema succinifaciens*, which has been uniquely identified in the gut microbiomes of pre-agricultural populations, as well as that of *Barnesiella intestinihominis*, which appears to be specifically present in the gut microbiomes of humans from developed countries.

Reconstruction of the *T. succinifaciens* genome was carried out using the Matses datasets due to its higher relative abundance in this population compared with that of Hadza datasets, and resulted in 97 contigs with a total length of 3.12 Mbp. Danish datasets were employed to reconstruct the *B. intestinihominis* genome, which generated 326 contigs encompassing 3.01 Mbp. *In silico* prediction of the glycobiome and identification of the pathways for degradation of carbohydrates and amino acids by means of the CAZy and MetaCyc databases provide detailed insights into the carbohydrate degradation capabilities of these two species (Fig. 4a).

Notably, glycobiome prediction revealed that *T. succinifaciens* and *B. intestinihominis* harbor 47 and 118 GH-encoding genes respectively (Fig. 4a). Furthermore, profiling of pathways for amino acid and carbohydrate degradation predicts that *B. intestinihominis* possesses a larger number of genes and pathways associated with the degradation of carbohydrates and amino acids, as well as with the metabolism of fatty acids and lipids as compared with *T. succinifaciens* (Fig. 4b).

These findings appear to reflect dietary differences between pre-agricultural and urbanized-industrialized communities (Cordain *et al.*, 2005). In fact, *B. intestinihominis* has only been identified in urbanized populations with a diet characterized by an increased diversity in polysaccharides and an augmented richness in proteins and lipids as compared with the Matses and Hadza pre-agricultural communities, which have been reported to follow a very simple (i.e. undiversified) diet (Cordain *et al.*, 2005). Functional data sets collected from human populations living in different countries across the globe highlight for the first time that changes in lifestyle and diet as caused by urbanization/industrialization appears to have created a selective pressure towards an expansion of the microbiome by acquisition of genes responsible for degradation and metabolism of a wider range of polysaccharides. In addition, an increase in protein and lipid intake also seems to have caused a marked expansion of the genetic repertoire involved in the metabolism of these energy sources.

Conclusions

In this mini-review, we explored the possibility that urbanization/industrialization processes have shaped the gut microbiomes as determined by a meta-analysis of various metagenomic datasets obtained from fecal samples of healthy human adults living in different countries across the world, including two pre-agricultural communities. There has been a considerable paucity of information regarding this topic. In fact, despite the increasing amount of scientific literature concerning the characterization of human gut microbiomes, very few are based on shotgun metagenomic sequencing focused on specific communities, while they typically do not involve analysis of datasets obtained from different geographical regions (Qin *et al.*, 2012; Karlsson *et al.*, 2013; Lim *et al.*, 2014; Feng *et al.*, 2015; Voigt *et al.*, 2015; Raymond *et al.*, 2016). The reconstructed WGMD allows the prediction of the global human gut core-microbiome, which was shown to encompass 22 genera present in the 18 datasets analysed, and representative of 15 nations. The comparison between pre-agricultural vs. urbanized/industrialized gut microbiomes of WGMD allowed the identification of particular taxa that seem to have been acquired (*Barnesiella intestinihominis*) or lost (*Treponema succinifaciens*) during the urbanization/industrialization process, perhaps as a result of dietary changes. In this context, the increase in dietary diversity and in protein and lipid intake appears to have caused an expansion of the metabolic capabilities of the human gut microbiome towards degradation of a wider range of polysaccharides and utilization of a higher number of available amino acids and lipids. However, the lack of metadata regarding dietary habits and lifestyle of the assessed populations prevents us to make any connection between diet and microbiome composition. Furthermore, antibiotic resistance profiling in the analysed datasets underlined a progressive increase in AREs proportional to the level of

urbanization/industrialization and corresponding intensity of antibiotic treatment.

As demonstrated in this mini-review, development and integration of worldwide metagenomic databases, e.g. the WGMD, will be pivotal for comprehensive studies combining novel metagenomic datasets and previous data from literature. In this context, meta-analyses of metagenomic datasets will aim to overcome the extensive fragmentation that still characterizes metagenomic studies, thus resulting in a more comprehensive overview of this rapidly expanding research area.

Acknowledgements

This work was funded by the EU Joint Programming Initiative – A Healthy Diet for a Healthy Life (JPI HDHL, http://www.healthydietforhealthylife.eu/) to MV and DvS (Grant no. 15/JP/HDHL/3280), and the MIUR to MV. We thank GenProbio srl for financial support of the Laboratory of Probiogenomics. LM is supported by Fondazione Cariparma, Parma, Italy. DvS is a member of The APC Microbiome Institute funded by Science Foundation Ireland (SFI), through the Irish Government's National Development Plan (Grant no. SFI/12/RC/2273). The authors declare that there is no conflict of interest for this study.

References

Brownlee, I.A., Allen, A., Pearson, J.P., Dettmar, P.W., Havler, M.E., Atherton, M.R., and Onsoyen, E. (2005) Alginate as a source of dietary fiber. *Crit Rev Food Sci Nutr* **45:** 497–510.

Burges Watson, D. (2008) Public health and carrageenan regulation: a review and analysis. *J Appl Phycol* **20:** 505–513.

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.

Caspi, R., Altman, T., Dreher, K., Fulcher, C.A., Subhraveti, P., Keseler, I.M., *et al.* (2012) The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. *Nucleic Acids Res* **40:** D742–D753.

Clemente, J.C., Pehrsson, E.C., Blaser, M.J., Sandhu, K., Gao, Z., Wang, B., *et al.* (2015) The microbiome of uncontacted Amerindians. *Sci Adv* **1:** pii e1500183.

Conlon, M.A., and Bird, A.R. (2015) The impact of diet and lifestyle on gut microbiota and human health. *Nutrients* **7:** 17–44.

Cordain, L., Eaton, S.B., Sebastian, A., Mann, N., Lindeberg, S., Watkins, B.A., *et al.* (2005) Origins and evolution of the Western diet: health implications for the 21st century. *Am J Clin Nutr* **81:** 341–354.

De Filippo, C., Cavalieri, D., Di Paola, M., Ramazzotti, M., Poullet, J.B., Massart, S., *et al.* (2010) Impact of diet in shaping gut microbiota revealed by a comparative study in children from Europe and rural Africa. *Proc Natl Acad Sci U S A* **107:** 14691–14696.

Dehingia, M., Devi, K.T., Talukdar, N.C., Talukdar, R., Reddy, N., Mande, S.S., *et al.* (2015) Gut bacterial diversity of the tribes of India and comparison with the worldwide data. *Sci Rep* **5:** 18563.

Feng, Q., Liang, S., Jia, H., Stadlmayr, A., Tang, L., Lan, Z., *et al.* (2015) Gut microbiome development along the colorectal adenoma-carcinoma sequence. *Nat Commun* **6:** 6528.

Giacani, L., and Lukehart, S.A. (2014) The endemic treponematoses. *Clin Microbiol Rev* **27:** 89–115.

Karlsson, F.H., Tremaroli, V., Nookaew, I., Bergstrom, G., Behre, C.J., Fagerberg, B., *et al.* (2013) Gut metagenome in European women with normal, impaired and diabetic glucose control. *Nature* **498:** 99–103.

Leung, M.H., Wilkins, D., and Lee, P.K. (2016) Erratum: Insights into the pan-microbiome: skin microbial communities of Chinese individuals differ from other racial groups. *Sci Rep* **6:** 21355.

Li, J., Jia, H., Cai, X., Zhong, H., Feng, Q., Sunagawa, S., *et al.* (2014) An integrated catalog of reference genes in the human gut microbiome. *Nat Biotechnol* **32:** 834–841.

Lim, M.Y., Rho, M., Song, Y.M., Lee, K., Sung, J., and Ko, G. (2014) Stability of gut enterotypes in Korean monozygotic twins and their association with biomarkers and diet. *Sci Rep* **4:** 7348.

Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P.M., and Henrissat, B. (2014) The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res* **42:** D490–D495.

Lugli, G.A., Milani, C., Mancabelli, L., van Sinderen, D., and Ventura, M. (2016) MEGAnnotator: a user-friendly pipeline for microbial genomes assembly and annotation. *FEMS Microbiol Lett* **363:** pii fnw049.

Martinez, I., Stegen, J.C., Maldonado-Gomez, M.X., Eren, A.M., Siba, P.M., Greenhill, A.R., and Walter, J. (2015) The gut microbiota of rural papua new guineans: composition, diversity patterns, and ecological processes. *Cell Rep* **11:** 527–538.

McArthur, A.G., Waglechner, N., Nizam, F., Yan, A., Azad, M.A., Baylay, A.J., *et al.* (2013) The comprehensive antibiotic resistance database. *Antimicrob Agents Chemother* **57:** 3348–3357.

Milani, C., Hevia, A., Foroni, E., Duranti, S., Turroni, F., Lugli, G.A., *et al.* (2013) Assessing the fecal microbiota: an optimized ion torrent 16S rRNA gene-based analysis protocol. *PLoS One* **8:** e68739.

Nishijima, S., Suda, W., Oshima, K., Kim, S.W., Hirose, Y., Morita, H., and Hattori, M. (2016) The gut microbiome of healthy Japanese and its microbial and functional uniqueness. *DNA Res* **23:** 125–133.

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* **6:** 6505.

Popkin, B.M., Adair, L.S., and Ng, S.W. (2012) Global nutrition transition and the pandemic of obesity in developing countries. *Nutr Rev* **70:** 3–21.

Prajapati, V.D., Jani, G.K., Zala, B.S., and Khutliwala, T.A. (2013) An insight into the emerging exopolysaccharide gellan gum as a novel polymer. *Carbohydr Polym* **93:** 670–678.

Qin, J., Li, R., Raes, J., Arumugam, M., Burgdorf, K.S., Manichanh, C., *et al.* (2010) A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* **464:** 59–65.

Qin, J., Li, Y., Cai, Z., Li, S., Zhu, J., Zhang, F., *et al.* (2012) A metagenome-wide association study of gut microbiota in type 2 diabetes. *Nature* **490:** 55–60.

Quail, M.A., Smith, M., Coupland, P., Otto, T.D., Harris, S.R., Connor, T.R., *et al.* (2012) A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers. *BMC Genomics* **13:** 341.

Rampelli, S., Schnorr, S.L., Consolandi, C., Turroni, S., Severgnini, M., Peano, C., *et al.* (2015) Metagenome sequencing of the Hadza Hunter-gatherer gut microbiota. *Curr Biol* **25:** 1682–1693.

Raymond, F., Ouameur, A.A., Deraspe, M., Iqbal, N., Gingras, H., Dridi, B., *et al.* (2016) The initial state of the human gut microbiome determines its reshaping by antibiotics. *ISME J* **10:** 707–720.

Round, J.L., and Mazmanian, S.K. (2009) The gut microbiota shapes intestinal immune responses during health and disease. *Nat Rev Immunol* **9:** 313–323.

Salonen, A., Salojarvi, J., Lahti, L., and de Vos, W.M. (2012) The adult intestinal core microbiota is determined by analysis depth and health status. *Clin Microbiol Infect* **18(Suppl 4):** 16–20.

Samuel, B.S., Hansen, E.E., Manchester, J.K., Coutinho, P.M., Henrissat, B., Fulton, R., *et al.* (2007) Genomic and metabolic adaptations of Methanobrevibacter smithii to the human gut. *Proc Natl Acad Sci U S A* **104:** 10643–10648.

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* **5:** 3654.

Truong, D.T., Franzosa, E.A., Tickle, T.L., Scholz, M., Weingart, G., Pasolli, E., *et al.* (2015) MetaPhlAn2 for enhanced metagenomic taxonomic profiling. *Nat Methods* **12:** 902–903.

Voigt, A.Y., Costea, P.I., Kultima, J.R., Li, S.S., Zeller, G., Sunagawa, S., and Bork, P. (2015) Temporal and technical variability of human gut metagenomes. *Genome Biol* **16:** 73.

Watanabe, Y., Nagai, F., and Morotomi, M. (2012) Characterization of Phascolarctobacterium succinatutens sp. nov., an asaccharolytic, succinate-utilizing bacterium isolated from human feces. *Appl Environ Microbiol* **78:** 511–518.

Wu, G.D., Chen, J., Hoffmann, C., Bittinger, K., Chen, Y.Y., Keilbaugh, S.A., *et al.* (2011) Linking long-term dietary patterns with gut microbial enterotypes. *Science* **334:** 105–108.

Xu, J., and Gordon, J.I. (2003) Honor thy symbionts. *Proc Natl Acad Sci U S A* **100:** 10452–10459.

Yatsunenko, T., Rey, F.E., Manary, M.J., Trehan, I., Dominguez-Bello, M.G., Contreras, M., *et al.* (2012) Human gut microbiome viewed across age and geography. *Nature* **486:** 222–227.

Zeller, G., Tap, J., Voigt, A.Y., Sunagawa, S., Kultima, J.R., Costea, P.I., *et al.* (2014) Potential of fecal microbiota for early-stage detection of colorectal cancer. *Mol Syst Biol* **10:** 766.

Zeng, X., and Lin, J. (2013) Beta-lactamase induction and cell wall metabolism in Gram-negative bacteria. *Front Microbiol* **4:** 128.

Supporting information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Fig. S1. Microbial profiling of the 142 samples at phylum level. Sample names and origins are reported in Supporting Information Tables S1.

Fig. S2. Taxonomic profiling of the 142 analyzed samples at genus level. Sample names and origins are reported in Supporting Information Tables S1.

Table S1. Summary of the observations discussed in this study, accompained by a comparison with previously published data

Table S2. List of publicly available data used in this study.

Table S3. Diet of pre-agricultural communities.

Table S4. List of pathway involved in carbohydrate degradation detected in analysis.

Chapter 6

Unveiling the gut microbiota composition and functionality associated with constipation through metagenomic analyses

Mancabelli L, Milani C, Lugli GA, Turroni F, Mangifesta M, Viappiani A, Ticinesi A, Nouvenne A, Meschi T, van Sinderen D, Ventura M.

This chapter was published in:

Scientific Reports. 2017 Aug 29;7(1):9879.

SCIENTIFIC REPORTS

OPEN

Unveiling the gut microbiota composition and functionality associated with constipation through metagenomic analyses

Received: 24 May 2017
Accepted: 9 August 2017
Published online: 29 August 2017

Leonardo Mancabelli[1], Christian Milani[1], Gabriele Andrea Lugli[1], Francesca Turroni[1], Marta Mangifesta[1,2], Alice Viappiani[2], Andrea Ticinesi[3], Antonio Nouvenne[3], Tiziana Meschi[3], Douwe van Sinderen[4] & Marco Ventura[1]

Functional constipation (FC) is a gastrointestinal disorder with a high prevalence among the general population. The precise causes of FC are still unknown and are most likely multifactorial. Growing evidence indicates that alterations of gut microbiota composition contribute to constipation symptoms. Nevertheless, many discrepancies exist in literature and no clear link between FC and gut microbiota composition has as yet been identified. In this study, we performed 16 S rRNA-based microbial profiling analysis of 147 stool samples from 68 FC individuals and compared their microbial profiles with those of 79 healthy subjects (HS). Notably, the gut microbiota of FC individuals was shown to be depleted of members belonging to *Bacteroides*, *Roseburia* and *Coprococcus 3*. Furthermore, the metabolic capabilities of the gut microbiomes of five FC and five HS individuals were evaluated through shotgun metagenomics using a MiSeq platform, indicating that HS are enriched in pathways involved in carbohydrate, fatty acid and lipid metabolism as compared to FC. In contrast, the microbiomes corresponding to FC were shown to exhibit high abundance of genes involved in hydrogen production, methanogenesis and glycerol degradation. The identified differences in bacterial composition and metabolic capabilities may play an important role in development of FC symptoms.

The human gastrointestinal tract is colonized by complex communities of microorganisms, i.e. the gut microbiota, that are involved in several physiological functions of the host. These encompass metabolic, nutritional, physiological and immunological processes that are vital to maintain the host's health status[1,2]. In this context, alterations in the gut microbiota composition have been linked to certain common human intestinal diseases, such as pseudomembranous colitis (CDI)[3–8], ulcerative colitis (UC)[9–11] and Crohn's disease (CD)[9,12,13]. However, changes in the gut microbiota composition are also considered to play a crucial role in the establishment of gut related disorders such as irritable bowel syndrome (IBS)[14–16]. Functional constipation (FC) is a common gastrointestinal disorder with a prevalence between 5% and 20% of the general population[17,18], provoking a significant impact on quality of life[19]. In fact, it can result in discomforts such as abdominal distension, abdominal pain, headache, dizziness and loss of appetite[20]. Despite its high prevalence, only a small number of studies have investigated its possible correlation with particular gut microbiota alterations. Additionally, most of these studies relied on culture-based methods that are unable to assess the unculturable portion of the gut microbiota[21,22]. Moreover, the functional implications of these alterations and their impact on host physiology have never been assessed. Recently, two metagenomic studies compared the microbial population of stools collected from constipated and healthy individuals, highlighting an altered fecal microbiome associated with constipation[23,24]. However, these studies were limited by the small sample size and heterogeneity of participants (including women or obese children). Furthermore, it is worth mentioning that the observed differences in microbiota composition among healthy and constipated patients suffer from a number of discrepancies[20].

[1]Laboratory of Probiogenomics, Department of Chemistry, Life Sciences, and Environmental Sustainability, University of Parma, Parma, Italy. [2]GenProbio srl, Parma, Italy. [3]Department of Medicine and Surgery, University of Parma, Italy; Dipartimento Medico-Geriatrico-Riabilitativo, Azienda Ospedaliero-Universitaria di Parma, Parma, Italy. [4]APC Microbiome Institute and School of Microbiology, Bioscience Institute, National University of Ireland, Cork, Ireland. Correspondence and requests for materials should be addressed to M.V. (email: marco.ventura@unipr.it)

Here, in order to identify a statistically significant and comprehensive correlation between microbiota and constipation, we performed 16S rRNA-based profiling analysis of 147 stool samples collected from 68 functional constipated (FC) and 79 healthy subjects (HS). Furthermore, in order to better understand the role of the microbiome and its metabolic impact on the host, the gut microbiome of a random subsampling of 10 samples, five FC and five HS samples, was reconstructed and analyzed in detail by shotgun metagenomic analyses.

Results and Discussion

Patient enrollment and collection of fecal samples. In this study, we collected and analyzed 147 human stool samples from Italian subjects. More specifically, we obtained 68 samples from individuals affected by functional constipation (FC), while 44 samples were collected from healthy subjects (HS). The HS dataset was supplemented with data from 35 samples that we published previously[8,25]. Notably, these samples had been collected and processed using the same protocols as followed for the 112 samples sequenced in the current study (see below for details). Moreover, analysis of variance of beta-diversity was performed between each pool of samples processed in different sequencing runs, and the obtained data showed absence of any batch effects (Fig. S1a). In order to identify microbial biomarkers of functional constipation across all age ranges, we selected individuals with an age ranging from 4 to 94 (average age: 42 ± 22 years) (Table S1). Remarkably, beta-diversity and PERMANOVA analyses displayed absence of age-related clustering of the samples (Fig. S1b). Moreover, the enrolled individuals were not taking prebiotics and/or probiotics, not undergoing antibiotic treatment or any other medical therapy (including those specific for functional constipation such laxatives for one week prior sampling) and not suffering from acute or severe intestinal diseases such as ulcerative colitis (UC), Crohn's disease, acute inflammatory bowel disease (IBD), intestinal cancer and enteritis. Notably, functionally constipated individuals also reporting symptoms typical of IBS-C, such as abdominal pain, were excluded from this study. The selected individuals affected by functional constipation fulfill the ROME-III criteria and manifested infrequent bowel movements that are defined as three or less defecations per week[26]. Notably, statistical assessment of diet homogeneity of FC and HS groups revealed absence of statistically significant differences (Table S2). In order to avoid discrepancies in the *in silico* data, all newly sequenced as well as previously published datasets included in this study were subjected to bioinformatic analysis using the same pipeline based on a custom script for the Qiime software suite and the same 16S rRNA database (see Methods for details).

Intra- and Inter-individual variability among healthy and functionally constipated subjects. Stool samples from the 147 individuals enrolled in this study were obtained in order to assess the microbiota composition based on 16S rRNA-based sequencing analysis, as described previously[27]. MiSeq-mediated sequencing of the samples produced a total of 18,673,728 reads with an average of 127,032 ± 69,090 reads per sample (Table S2). Quality and chimera filtering produced a total of 10,164,847 filtered reads with an average of 69,149 filtered reads per sample, and ranging from 185,347 to 10,440 reads (Table S2).

Evaluation of rarefaction curves obtained through the Shannon and Chao1 biodiversity indeces calculated for 10 sub-samplings of sequenced read pools showed that both curves tend to reach a plateau. Therefore, in all cases the retrieved sequencing data is considered adequate to cover the vast majority of biodiversity contained within the samples (Fig. 1). Interestingly, average rarefaction curves revealed a difference between FC and HS samples in that, on average, the former samples were shown to exhibit a higher level of gut microbiota complexity compared to the latter samples. Statistical analysis, calculated for the highest sub-sampling point reached by all samples, i.e. 30,000 reads, showed that the two curves significantly differ based on a one-way analysis of variance (ANOVA) (p-value < 0.05). The observed dissimilarity of the alpha-diversity between FC and HS is in accordance with a previous study[24], being indicative of differences in bacterial composition and corresponding metabolic potential (see below). In order to evaluate the inter-individual differences between FC and HS samples in more detail, we assessed the beta-diversity[28] by means of unweighted UniFrac[29] and represented the results through a 3-Dimensional Principal Coordinate Analysis (PCoA). The PCoA plot shows that the majority of the samples are grouped as two different clusters that correspond to FC or HS individuals, thus highlighting an intriguing difference in microbiota composition between individuals that suffer from functional constipation and healthy individuals (Fig. 1). Notably, the obtained results were statistically supported by PERMANOVA analyses (p-value of <0.001).

Taxonomic profiling of the gut microbiota of functionally constipated and healthy individuals. Inspection of the predicted taxonomic profiles showed that both FC and HS samples possess a preponderant presence of the phyla Bacteroidetes (50.44% ± 18.29% and 58.37% ± 16.59%, respectively; p-value < 0.05) and Firmicutes (44.19% ± 17.02% and 36.65% ± 15.65%, respectively; p-value < 0.05), although with a significantly different average relative abundance (Fig. S2). When analyzed at genus level, the FC group displayed high levels of *Bacteroides* (34.25% ± 18.56%), *Faecalibacterium* (6.85% ± 6.19%), *Alistipes* (6.48% ± 9.69%), *Lachnospira* (4.44% ± 6.13%) and Unclassified member of Lachnospiraceae family (3.92% ± 2.40%). Similarly, the most represented taxa detected in HS samples were *Bacteroides* (45.23% ± 17.90%), *Alistipes* (5.34% ± 5.60%) and Unclassified member of Lachnospiraceae family (4.66% ± 4.38%).

Comparative analysis of the 331 bacterial taxa predicted by genus-level analysis revealed that 23 genera appeared to be present only in HS samples, while 17 were uniquely present in FC subjects (Table. S3). Analysis of the proportion of these unique genera found in each group, i.e. the prevalence, showed that these taxa are present in <20% of FC or HS profiles, thus indicating the absence of specific microbial biomarkers whose presence or absence is associated with constipation while pointing at a probable role played by the overall gut microbiota at functional level.

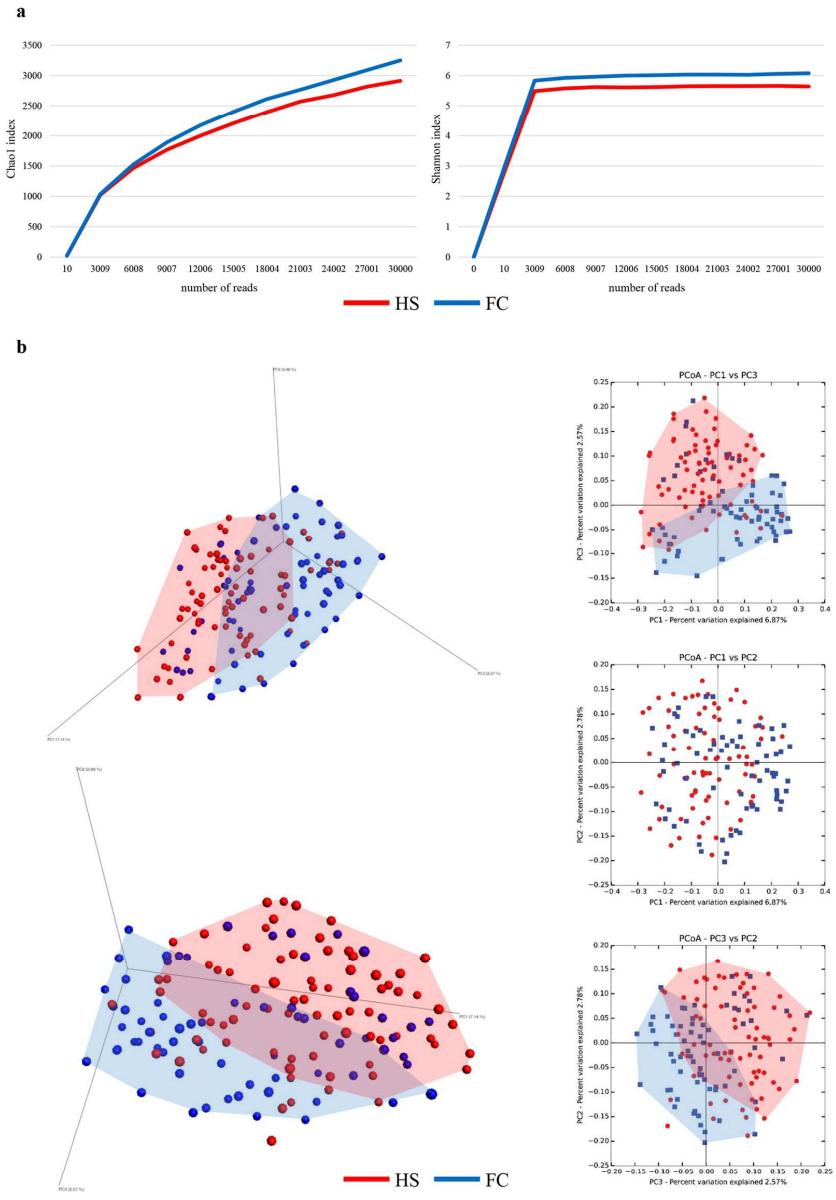
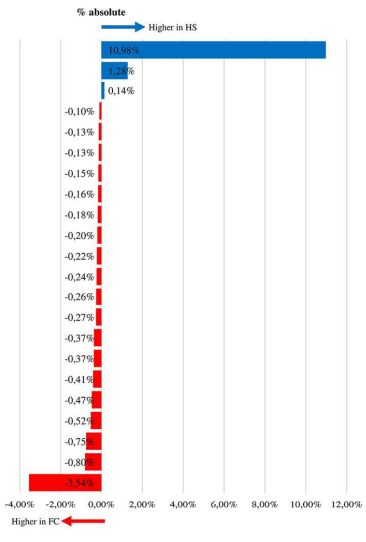


Figure 1. Evaluation of the alpha- and beta-diversity in the 147 analyzed samples. Panel a reports the average rarefaction curves based on the Chao1 and Shannon index at increasing sequencing depth of FC and HS samples. Panel b shows the predicted PCoA through two three-dimensional images and two-dimensional sections. FC and HS datasets and corresponding clusters are colored in blue and red, respectively.

Difference in gut microbiota composition. In order to evaluate possible differences in bacterial composition, ANOVA statistical analysis was employed to compare the average relative abundance in FC and HS groups of genera with an absolute percentage difference >0.1% (Fig. 2). Interestingly, the comparison between HS and FC datasets showed that profiles obtained from HS individuals are characterized by a statistically significant over-representation of *Bacteroides* (% absolute 1.28%, p-value < 0.01), *Roseburia* (% absolute 1.28%, p-value < 0.01) and *Coprococcus 3* (% absolute 0.14%, p-value < 0.01) and a statistically significant



Taxonomy	Relative abundance: Healthy (HS)	Relative abundance: Functional constipation (FC)	% relative	Prevalence: Healthy (HS)	Prevalence: Functional constipation (FC)	P-value
Bacteroides	45.23%	34.25%	32.04%	100.00%	100.00%	0.0004
Roseburia	3.24%	1.97%	64.91%	100.00%	100.00%	0.0060
Coprococcus 3	0.19%	0.05%	296.02%	96.20%	88.24%	0.0033
Ruminococcaceae UCG-003	0.12%	0.22%	-44.06%	67.09%	77.94%	0.0409
Butyricimonas	0.20%	0.33%	-38.58%	75.95%	79.41%	0.0098
Ruminiclostridium 5	0.19%	0.32%	-41.14%	100.00%	100.00%	0.0094
Ruminiclostridium	0.09%	0.24%	-61.38%	97.47%	100.00%	0.0039
U. m. of *Bacteroidales* order	0.05%	0.21%	-74.32%	96.20%	100.00%	0.0038
U. m. of *Clostridiales* order	0.08%	0.26%	-69.33%	91.14%	100.00%	0.0008
U. m. of *Porphyromonadaceae* family	0.09%	0.29%	-68.39%	81.01%	94.12%	0.0005
Ruminococcaceae NK4A214 group	0.17%	0.40%	-56.08%	77.22%	80.88%	0.0012
U. m. of *NB1-n* order	0.00%	0.24%	-98.73%	20.25%	42.65%	0.0011
Paraprevotella	0.17%	0.43%	-59.69%	40.51%	47.06%	0.0320
U. m. of *Ruminococcaceae* family	0.32%	0.58%	-45.46%	100.00%	100.00%	0.0000
Ruminococcaceae UCG-010	0.09%	0.46%	-80.31%	72.15%	73.53%	0.0004
Ruminococcaceae UCG-005	0.43%	0.80%	-46.47%	89.87%	95.59%	0.0176
Ruminiclostridium 6	0.33%	0.74%	-55.06%	67.09%	73.53%	0.0345
U. m. of *Mollicutes* order	0.05%	0.52%	-90.75%	26.58%	55.88%	0.0320
Phascolarctobacterium	0.46%	0.98%	-53.20%	79.75%	75.00%	0.0032
U. m. of *Clostridiales vadinBB60* group family	0.25%	1.00%	-75.21%	84.81%	89.71%	0.0021
Ruminococcaceae UCG-014	0.41%	1.21%	-65.90%	82.28%	85.29%	0.0025
Faecalibacterium	3.31%	6.85%	-51.71%	98.73%	100.00%	0.0001

Figure 2. Exploration of the diversity in HS and FC groups. The bar plot reports only genera with an absolute percentage difference between HS and FC averages >0.1% and a p-value < 0.05, evaluated by means of ANOVA statistical analysis. The table indicates the bacterial genera, the relative abundance and the prevalence of each group, the relative percentage difference and the p-value.

under-representation of genera belonging to the Ruminococcaceae family such as *Faecalibacterium* (% absolute −3.54%, p-value < 0.01) (Fig. 2). The depletion of the *Bacteroides* genus in FC samples may be correlated with alterations of the intestinal motility and secretory functions due to changes in the amount of available physiologically active substances in the metabolic environment of the gut[20,30]. In fact, the higher abundance of butyrate-producing taxa, such as *Coprococcus* and *Roseburia*, observed in HS samples may explain a faster colonic transit due to the motility-stimulating effect exerted by butyrate in the gut[1]. In this context, previous studies have reported that butyrate-producing taxa may stimulate colonic motility by induction of serotonin release or by facilitating cholinergic pathways by means of butyrate production[31,32]. In contrast, despite being a butyrate-producer, *Faecalibacterium* is significantly more abundant in FC samples. Interestingly, this genus has been reported to contribute to the pathogenesis of constipation via several mechanisms, such as inhibition of mucin secretion and reduction of stool volume[24].

Metagenomic characterization of FC and HS microbiomes. A total of 10 individuals were selected among the two clusters obtained from the PCoA analysis (Fig. 1b) as representatives of the FC and HS groups, and total bacterial DNA extracted from corresponding fecal samples was subjected to Illumina shotgun sequencing. These samples were selected on the basis of their microbial profiles that were shown to be closer to the average for each group. Collected read pools ranged from 8,656,289 to 2,885,092 after quality filtering, with an average number of reads per sample of 6,130,802 (Table S4). These data were then utilized for the reconstruction of metabolic pathway profiles in the analyzed microbiomes by means of a custom script based on the MetaCyc database[33]. A comparison between the averages of functionally constipated and the averages of healthy samples showed a significant difference in 629 pathways (p-value < 0.05). Of the latter, 327 and 302 pathways were more abundant in FC and HS samples, respectively (Table S5). Interestingly, the HS samples exhibited a higher abundance of genes (p-value < 0.05) involved in carbohydrate (increase of 21.15%) and fatty acid metabolism (increase of 25.93%) as compared to FC individuals (Fig. 3b). These pathways are implicated in production of short-chain fatty acids (SCFAs) and may play a role in stimulating ileal propulsive contractions through an enteric cholinergic reflex, thereby counteracting functional constipation[20,34–36]. Moreover, FC samples were shown to contain a higher abundance of genes involved in methanogenic pathways (increase of 24.96%, p-value < 0.01) and a predicted higher capability to produce hydrogen (increase of 113.69%, p-value 0.05) as compared to HS (Table S5, Fig. 3b). Thus, our data are consistent with previous observations and suggest that the (abundant) presence of methanogenic and H_2-consuming populations influence colonic motility and visceral sensitivity, and generate chronic constipation along with several correlated symptoms, such as flatulence and abdominal distension[37,38]. Interestingly, datasets obtained from functionally constipated individuals showed lower abundance of genes involved in methylglyoxal degradation as compared to healthy samples (decrease of −25.18%, p-value < 0.05) (Fig. 3b). Methylglyoxal is produced by intestinal bacteria and it is reported be a potential toxic metabolite that can be involved in many gut diseases, including functional constipation[39]. Furthermore, comparison between HS and FC samples highlights a significant difference in pathways implicated in

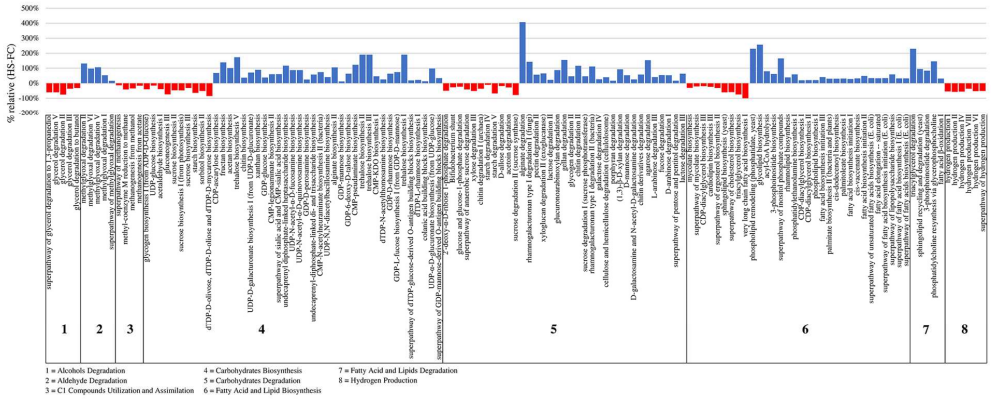


Figure 3. Functional characterization of FC and HS microbiomes. The bar plot shows the relative difference between the average abundance in HS and FC subjects of each pathway displaying ANOVA with p-value < 0.05.

glycerol degradation (decrease of −58.93% in HS, p-value < 0.05) (Fig. 3b). Glycerol is known to cause an osmotic effect in the rectum and for this reason is used to treat constipation[40]. Therefore, the increased ability to degrade glycerol by FC samples may induce and promote the symptoms of constipation. Nevertheless, despite the fact that the use of laxatives in the week before sampling was considered as a exclusion criterion, this observation may be linked to previous and prolonged use of PEG or glycerol suppositories.

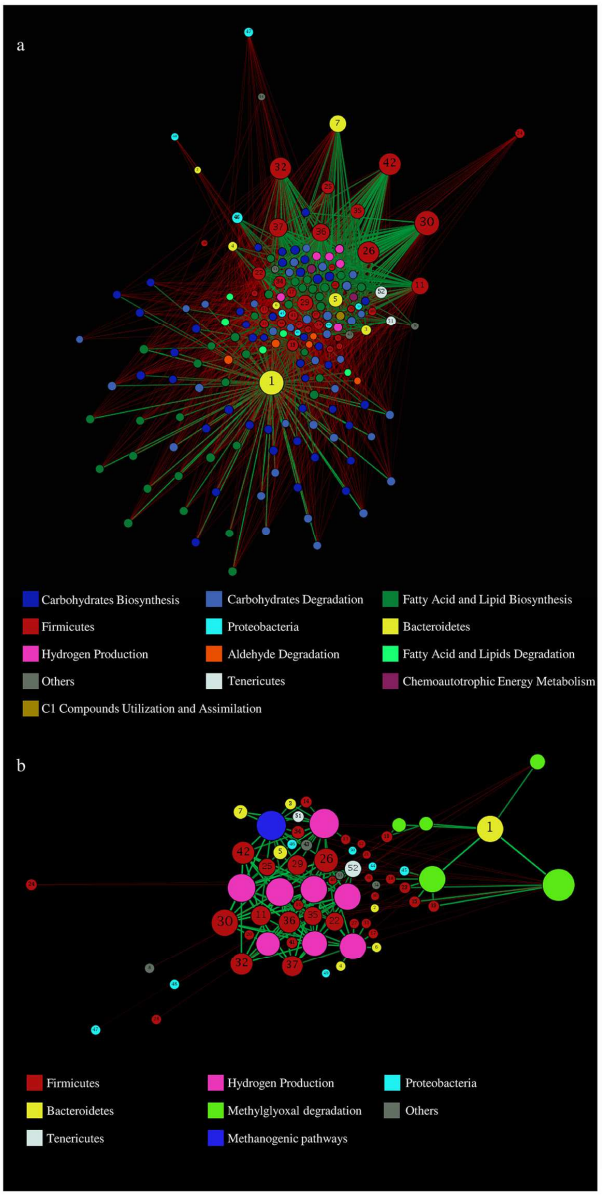
In addition, we evaluated co-variance between the bacterial genera and metabolic pathways displaying difference in abundance (ANOVA p-value < 0.05) in FC as compared to HS subjects. Force-driven network representation of these data revealed that *Bacteroides* (124 co-variances, p-value < 0.05), genera belonging Ruminococcaceae family (a total of 270 co-variances, p-value < 0.05) and the methanogenic genus *Gelria* (108 interactions, p-value < 0.05) exert a key role in modulating the metabolic functionalities of the gut microbiome that are altered during FC (Fig. 4a).

Moreover, a force-driven network representation based on the above-discussed pathways that may be related to FC development revealed that the Firmicutes phylum positively correlates with pathways involved in hydrogen production and methanogenesis, while *Bacteroides* correlates positively with the methylglyoxal degradation pathways (Fig. 4b). These results reflect the putative beneficial role played by *Bacteroides* to counter FC as also suggested by the higher abundance of this genus in HS with respect to FC samples (see above). Notably, an in depth analysis of 16S rRNA OTUs of HS samples revealed that 17.33% of the reads belonging to the *Bacteroides* genus correspond to unknown species, followed by *Bacteroides vulgatus* (13.83%), *Bacteroides uniformis* (4.21%) and *Bacteroides fragilis* (1.74%), thus indicating that further genomic analyses are still needed to shed light on the biological role of this genus in protecting against or in preventing FC.

Conclusions

Functional constipation is a widespread gastrointestinal disorder responsible for difficult or infrequent bowel movements defined as three or less defecations per week[26]. Despite the high worldwide prevalence of FC, a clear anatomical or physiological cause for this disorder has yet to be identified, thus pointing at a possible role exerted by the gut microbiota. Here, we confirm preliminary findings regarding gut microbiota compositional shifts in individuals affected by FC as compared to healthy controls, which are concurrent with a statistically significant increase of the gut microbiota biodiversity. Moreover, statistical analysis revealed alterations in relative abundance of specific taxa, such as *Bacteroides* and *Feacalibacterium*. Disregarding the fact that the cause of such taxonomic changes can't be linked to a specific physiological cause and may simply reflect altered transit time or diet, these taxa will be pivotal for diagnostic and prophylactic purposes as statistically-supported microbial biomarkers of constipation. Intriguingly, identification of taxa typically associated to a healthy gut status, e.g. the anti-inflammatory genus *Feacalibacterium*, as a biomarker of functional constipation, highlights that increased relative abundance of certain taxa in the presence of a gut disorder may not reflect a taxa-specific role in pathogenesis but may be linked to a global alteration of gut microbiota's homeostasis. Notably, the taxonomic profiles retrieved from faecal samples include both the autochthonous as well as the allochthonous microbiota, and may thus not be fully representative of the resident gut microbial population. In contrast, the use of biopsies from mucosal samples would be ideal to provide information on the indigenous microbiota[41]. Nevertheless, collection of fecal samples is less invasive and does not require specific clinical procedures. Thus, for a rapid screening aimed at the identification of biomarkers associated with specific disorders, e.g. functional constipation, the use of stool samples does not suffer from the aforementioned problems associated with collection of biopsies.

While exploration of the functional role of the gut microbiota in fecal transit time[42] has been attempted in the past through analysis of urine metabolites, the whole gut microbiome metabolic potential and its impact on host physiology and development of functional constipation has yet to be elucidated. To overcome this gap, we profiled metabolic pathways of microbiomes corresponding to functionally constipated and healthy individuals.



Id	Taxonomy
1	Bacteroides
2	Butyricimonas
3	U. m. of Porphyromonadaceae family
4	Paraprevotella
5	U. m. of Bacteroidales order
6	Flavobacterium
7	U. m. of Bacteroidetes phylum
8	Cercis gigantea
10	Catabacter
11	U. m. of Clostridiales vadinBB60 group family
12	Defluviitaleaceae UCG-011
13	Family XIII AD3011 group
14	Coprococcus 2
15	Coprococcus 3
16	Lachnoanaerobaculum
17	Lachnospiraceae NC2004 group
18	Lachnospiraceae NK4B4 group
19	Roseburia
20	Dehalobacterium
21	Peptococcus
22	U. m. of Peptococcaceae family
23	U. m. of Peptostreptococcaceae family
24	Faecalibacterium
25	Hydrogenoanaerobacterium
26	Papillibacter
27	Ruminiclostridium 5
28	Ruminiclostridium 6
29	Ruminiclostridium
30	Ruminococcaceae NK4A214 group
31	Ruminococcaceae UCG-004
32	Ruminococcaceae UCG-005
33	Ruminococcaceae UCG-007
34	Ruminococcaceae UCG-010
35	Ruminococcaceae UCG-014
36	U. m. of Ruminococcaceae family
37	U. m. of Clostridiales order
38	Erysipelatoclostridium
40	Lactococcus
41	Phascolarctobacterium
42	Gelria
43	Victivallis
44	Lautropia
45	Polaromonas
46	Oxalobacter
47	Helicobacter
48	Desulfovibrio
49	U. m. of Rhodospirillaceae family
50	U. m. of Proteobacteria phylum
51	U. m. of Mollicutes order
52	U. m. of NB1-n order
53	U. m. of Verrucomicrobia phylum
54	U. m. of Bacteria kingdom

Figure 4. Co-variance network of bacterial genera and metabolic pathways with statistically significant difference in abundance between FC and HS subjects. Panel a shows a force-driven network based on the predicted co-variances with p-values < 0.05 between the genera and metabolic pathways identified as statistically altered in FC as compared to HS. Panel b reports a force-driven network based on the predicted co-variances with p-values < 0.05 between genera and metabolic pathways putatively involved in FC development. Co-variances with p-values < 0.05 are used to define the attractive or repulsive force of the edges. The node size is proportional to the number of co-variances. Node color indicates phylum or pathway category as reported in the image. The taxonomy, i.e. bacterial genera, of the nodes is indicated with number, as listed in the table.

The here reconstructed gut microbiome of individuals affected by functional constipation revealed for the first time that the FC microbiome is characterized by a high abundance of genes involved in hydrogen production, methanogenesis and glycerol degradation. In contrast, the microbiomes of HS samples showed an increase of pathways implicated in carbohydrate and fatty acid metabolism, and in methylglyoxal degradation. Alteration of these metabolic pathways appears to impact on functional constipation and related symptoms, thus highlighting the key functional role exerted by the gut microbiome in maintaining the health status of the host. Nevertheless, due to the limited number of samples that were analysed by shotgun sequencing, additional experiments are needed to validate these observations.

Altogether, taxonomic and functional data reported in this study represent a solid base for future development of both prophylactic screenings and therapies for functional constipation based on alterations of gut microbiota composition through personalized diet or pre- and pro-biotic treatments.

Methods

Datasets included in this study. We enrolled 68 volunteers suffering from functional constipation (FC) and 44 healthy subjects (HS) in an outpatient clinic setting. The individuals affected by functional constipation presented infrequent bowel movements and fulfilled the ROME-III criteria. A stool sample, consisting of 6–10 g fresh fecal material, was obtained from each subject and immediately frozen at −80 °C until it was processed for DNA extraction. DNA was extracted from each stool sample using the QIAamp DNA Stool Mini kit (Qiagen Ltd, Strasse, Germany) following the manufacturer's instructions (Qiagen Ltd). The study protocol was approved by the Ethics Committee of the University of Parma. Informed consent was obtained from all participants. All investigations were carried out following the principles of the Declaration of Helsinki.

16 S rRNA gene amplification. Partial 16 S rRNA gene sequences were amplified from extracted DNA using primer pair Probio_Uni and/Probio_Rev, which targets the V3 region of the 16 S rRNA gene sequence[27]. Illumina adapter overhang nucleotide sequences were then added to the partial 16 S rRNA gene-specific amplicons, which in turn were further processed by employing the 16 S Metagenomic Sequencing Library Preparation Protocol (Part no. 15044223 Rev. B—Illumina; see also below). Amplifications were carried out using a Verity Thermocycler (Applied Biosystems). The integrity of the PCR amplicons was analyzed by electrophoresis on a 2200 TapeStation Instrument (Agilent Technologies, USA).

MiSeq sequencing of 16 S rRNA gene-based amplicons. PCR products obtained following amplification of part of the 16 S rRNA gene sequences were purified by a magnetic purification step involving Agencourt AMPure XP DNA purification beads (Beckman Coulter Genomics GmbH, Bernried, Germany) in order to remove primer dimers. The DNA concentration of the amplified sequence library was estimated through fluorimetric Qubit quantification system (Life Technologies). Amplicons were diluted to 4 nM and 5 µl of each diluted DNA amplicon sample was mixed to prepare the pooled final library. Sequencing was performed using an Illumina MiSeq sequencer with MiSeq Reagent Kit v3 chemicals.

Analysis of 16 S rRNA microbial profiling datasets. The fastq files were processed using QIIME[28] as previously described[27]. Paired-end reads were merged, and quality control implementation allowed the retention of sequences with a length between 140 and 400 bp, mean sequence quality score >25 and with truncation of a sequence at the first base if a low quality within a rolling 10-bp window was found. Sequences with mismatched forward and/or reverse primers were omitted. 16 S rRNA operational taxonomic units (OTUs) were defined at $\geq 97\%$ sequence homology using uclust[43]. All reads were classified to the lowest possible taxonomic rank using QIIME[28] and a reference dataset from the SILVA database v. 123[44]. The microbial richness of the samples (alpha-diversity) was evaluated with the Chao1 and Shannon index through the alpha_rarefaction.py script included in the Qiime software suite using default parameters. Similarities between samples (beta-diversity) were calculated by unweighted uniFrac[45]. Principal coordinate analysis (PCoA) representations of beta-diversity were performed using QIIME[28].

Shotgun metagenomics. DNA library preparation was performed using the Nextera XT DNA sample preparation kit (Illumina, San Diego, CA) according to the manufacturer's instructions. In brief, 1 ng input DNA from each sample was used for library preparation. The isolated DNA underwent fragmentation, adapter ligation and amplification. The ready-to-go libraries were pooled equimolarly, denaturated and diluted to a sequencing concentration of 1.8 pM. Sequencing was performed on NextSeq. 550 instrument (Illumina, San Diego, CA), according to the manufacturer's instructions, using the 2×150 bp High Output sequencing kit, and spike-in of 1% PhiX control library.

Analysis of metagenomic datasets. The generated paired fastq files were filtered for reads with a quality score of <25, for sequences of human DNA, as well as for reads <80 bp. Bases were also removed from the end of the reads unless the average quality score in a window of 5 bp was >25. Reconstruction of bacterial metabolic pathways was performed using custom scripts based on htseq-count[46] and the MetaCyc database[33], respectively.

Statistical analyses. QIIME and SPSS software (www.ibm.com/software/it/analytics/spss/) were used to complete statistical analysis. All data were presented as means ± SEM. PERMANOVA were performed using 999 permutations to estimate p-values for differences among populations. Furthermore, differential abundance of bacteria taxa and metabolic pathways were tested by one-way analysis of variance (ANOVA).

Data Deposition. The 16S rRNA profiling data sequenced in this study were deposited in the Sequence Read Archive (SRA) database under the SRP106879 accession number. Shotgun metagenomics data are accessible through SRA study accession number SRP106935.

References

1. Chassard, C. *et al.* Functional dysbiosis within the gut microbiota of patients with constipated-irritable bowel syndrome. *Alimentary pharmacology & therapeutics* **35**, 828–838, doi:10.1111/j.1365-2036.2012.05007.x (2012).
2. Round, J. L. & Mazmanian, S. K. The gut microbiota shapes intestinal immune responses during health and disease. *Nat Rev Immunol* **9**, 313–323, doi:10.1038/nri2515 (2009).
3. Zhang, L. *et al.* Insight into alteration of gut microbiota in Clostridium difficile infection and asymptomatic C. difficile colonization. *Anaerobe* **34**, 1–7, doi:10.1016/j.anaerobe.2015.03.008 (2015).
4. Rojo, D. *et al.* Clostridium difficile heterogeneously impacts intestinal community architecture but drives stable metabolome responses. *The ISME journal* **9**, 2206–2220, doi:10.1038/ismej.2015.32 (2015).
5. Khanna, S. *et al.* Gut microbiome predictors of treatment response and recurrence in primary Clostridium difficile infection. *Alimentary pharmacology & therapeutics* **44**, 715–727, doi:10.1111/apt.13750 (2016).
6. Abujamel, T. *et al.* Defining the vulnerable period for re-establishment of Clostridium difficile colonization after treatment of C. difficile infection with oral vancomycin or metronidazole. *PloS one* **8**, e76269, doi:10.1371/journal.pone.0076269 (2013).
7. Gu, S. *et al.* Identification of key taxa that favor intestinal colonization of Clostridium difficile in an adult Chinese population. *Microbes and infection* **18**, 30–38, doi:10.1016/j.micinf.2015.09.008 (2016).
8. Milani, C. *et al.* Gut microbiota composition and Clostridium difficile infection in hospitalized elderly individuals: a metagenomic study. *Scientific reports* **6**, 25945, doi:10.1038/srep25945 (2016).
9. Tong, M. *et al.* A modular organization of the human intestinal mucosal microbiota and its association with inflammatory bowel disease. *PloS one* **8**, e80702, doi:10.1371/journal.pone.0080702 (2013).
10. Shah, R. *et al.* Composition and function of the pediatric colonic mucosal microbiome in untreated patients with ulcerative colitis. *Gut microbes* **7**, 384–396, doi:10.1080/19490976.2016.1190073 (2016).
11. Duranti, S. *et al.* Elucidating the gut microbiome of ulcerative colitis: bifidobacteria as novel microbial biomarkers. *FEMS microbiology ecology* 92, doi:10.1093/femsec/fiw191 (2016).
12. Weir, T. L. *et al.* Stool microbiome and metabolome differences between colorectal cancer patients and healthy adults. *PloS one* **8**, e70803, doi:10.1371/journal.pone.0070803 (2013).
13. Mondot, S. *et al.* Structural robustness of the gut mucosal microbiota is associated with Crohn's disease remission after surgery. *Gut* **65**, 954–962, doi:10.1136/gutjnl-2015-309184 (2016).
14. Pozuelo, M. *et al.* Reduction of butyrate- and methane-producing microorganisms in patients with Irritable Bowel Syndrome. *Scientific reports* **5**, 12693, doi:10.1038/srep12693 (2015).
15. Saulnier, D. M. *et al.* Gastrointestinal microbiome signatures of pediatric patients with irritable bowel syndrome. *Gastroenterology* **141**, 1782–1791, doi:10.1053/j.gastro.2011.06.072 (2011).
16. Ringel-Kulka, T. *et al.* Molecular characterization of the intestinal microbiota in patients with and without abdominal bloating. *American journal of physiology. Gastrointestinal and liver physiology* **310**, G417–426, doi:10.1152/ajpgi.00044.2015 (2016).
17. Lovell, R. M. & Ford, A. C. Global prevalence of and risk factors for irritable bowel syndrome: a meta-analysis. *Clinical gastroenterology and hepatology: the official clinical practice journal of the American Gastroenterological Association* **10**, 712–721 e714, doi:10.1016/j.cgh.2012.02.029 (2012).
18. Suares, N. C. & Ford, A. C. Prevalence of, and risk factors for, chronic idiopathic constipation in the community: systematic review and meta-analysis. *The American journal of gastroenterology* 106, 1582-1591; quiz 1581, 1592, doi:10.1038/ajg.2011.164 (2011).
19. Belsey, J., Greenfield, S. & Candy, D. & Geraint, M. Systematic review: impact of constipation on quality of life in adults and children. *Alimentary pharmacology & therapeutics* **31**, 938–949, doi:10.1111/j.1365-2036.2010.04273.x (2010).
20. Zhao, Y. & Yu, Y. B. Intestinal microbiota and chronic constipation. *SpringerPlus* **5**, 1130, doi:10.1186/s40064-016-2821-1 (2016).
21. Khalif, I. L., Quigley, E. M., Konovitch, E. A. & Maximova, I. D. Alterations in the colonic flora and intestinal permeability and evidence of immune activation in chronic constipation. *Digestive and liver disease: official journal of the Italian Society of Gastroenterology and the Italian Association for the Study of the Liver* **37**, 838–849, doi:10.1016/j.dld.2005.06.008 (2005).
22. Zoppi, G. *et al.* The intestinal ecosystem in chronic functional constipation. *Acta paediatrica* **87**, 836–841 (1998).
23. Parthasarathy, G. *et al.* Relationship Between Microbiota of the Colonic Mucosa vs Feces and Symptoms, Colonic Transit, and Methane Production in Female Patients With Chronic Constipation. *Gastroenterology* **150**, 367–379 e361, doi:10.1053/j.gastro.2015.10.005 (2016).
24. Zhu, L. *et al.* Structural changes in the gut microbiome of constipated patients. *Physiological genomics* **46**, 679–686, doi:10.1152/physiolgenomics.00082.2014 (2014).
25. Ferrario, C. *et al.* Modulation of fecal Clostridiales bacteria and butyrate by probiotic intervention with Lactobacillus paracasei DG varies among healthy adults. *The Journal of nutrition* **144**, 1787–1796, doi:10.3945/jn.114.197723 (2014).
26. Jamshed, N., Lee, Z. E. & Olden, K. W. Diagnostic approach to chronic constipation in adults. *American family physician* **84**, 299–306 (2011).
27. Milani, C. *et al.* Assessing the fecal microbiota: an optimized ion torrent 16S rRNA gene-based analysis protocol. *PloS one* **8**, e68739, doi:10.1371/journal.pone.0068739 (2013).
28. Caporaso, J. G. *et al.* QIIME allows analysis of high-throughput community sequencing data. *Nature methods* **7**, 335–336, doi:10.1038/nmeth.f.303 (2010).
29. Lozupone, C., Lladser, M. E., Knights, D., Stombaugh, J. & Knight, R. UniFrac: an effective distance metric for microbial community comparison. *The ISME journal* **5**, 169–172, doi:10.1038/ismej.2010.133 (2011).
30. Kim, S. E. *et al.* Change of Fecal Flora and Effectiveness of the Short-term VSL#3 Probiotic Treatment in Patients With Functional Constipation. *Journal of neurogastroenterology and motility* **21**, 111–120, doi:10.5056/jnm14048 (2015).
31. Reigstad, C. S. *et al.* Gut microbes promote colonic serotonin production through an effect of short-chain fatty acids on enterochromaffin cells. *FASEB journal: official publication of the Federation of American Societies for Experimental Biology* **29**, 1395–1403, doi:10.1096/fj.14-259598 (2015).
32. Soret, R. *et al.* Short-chain fatty acids regulate the enteric neurons and control gastrointestinal motility in rats. *Gastroenterology* **138**, 1772–1782, doi:10.1053/j.gastro.2010.01.053 (2010).
33. Caspi, R. *et al.* The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of Pathway/Genome Databases. *Nucleic acids research* **42**, D459–471, doi:10.1093/nar/gkt1103 (2014).
34. Coffin, B. *et al.* Local regulation of ileal tone in healthy humans. *The American journal of physiology* **272**, G147–153 (1997).
35. Kamath, P. S., Phillips, S. F. & Zinsmeister, A. R. Short-chain fatty acids stimulate ileal motility in humans. *Gastroenterology* **95**, 1496–1502 (1988).
36. Ge, X. *et al.* Potential role of fecal microbiota from patients with slow transit constipation in the regulation of gastrointestinal motility. *Scientific reports* **7**, 441, doi:10.1038/s41598-017-00612-y (2017).

37. Attaluri, A., Jackson, M., Valestin, J. & Rao, S. S. Methanogenic flora is associated with altered colonic transit but not stool characteristics in constipation without IBS. *The American journal of gastroenterology* **105**, 1407–1411, doi:10.1038/ajg.2009.655 (2010).
38. Soares, A. C., Lederman, H. M., Fagundes-Neto, U. & de Morais, M. B. Breath hydrogen test after a bean meal demonstrates delayed oro-cecal transit time in children with chronic constipation. *Journal of pediatric gastroenterology and nutrition* **41**, 221–224 (2005).
39. Zhang, S. *et al.* Methylglyoxal induces systemic symptoms of irritable bowel syndrome. *PloS one* **9**, e105307, doi:10.1371/journal.pone.0105307 (2014).
40. Portalatin, M. & Winstead, N. Medical management of constipation. *Clinics in colon and rectal surgery* **25**, 12–19, doi:10.1055/s-0032-1301754 (2012).
41. Turroni, F., van Sinderen, D. & Ventura, M. Bifidobacteria: from ecology to genomics. *Frontiers in bioscience* **14**, 4673–4684 (2009).
42. Roager, H. M. *et al.* Colonic transit time is related to bacterial metabolism and mucosal turnover in the gut. *Nature microbiology* **1**, 16093, doi:10.1038/nmicrobiol.2016.93 (2016).
43. Edgar, R. C. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**, 2460–2461, doi:10.1093/bioinformatics/btq461 (2010).
44. Quast, C. *et al.* The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic acids research* **41**, D590–596, doi:10.1093/nar/gks1219 (2013).
45. Lozupone, C. & Knight, R. UniFrac: a new phylogenetic method for comparing microbial communities. *Appl Environ Microbiol* **71**, 8228–8235, doi:10.1128/AEM.71.12.8228-8235.2005 (2005).
46. Anders, S., Pyl, P. T. & Huber, W. HTSeq–a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166–169, doi:10.1093/bioinformatics/btu638 (2015).

Acknowledgements

This work was funded by the EU Joint Programming Initiative – A Healthy Diet for a Healthy Life (JPI HDHL, http://www.healthydietforhealthylife.eu/) to MV and DvS, and the MIUR to MV. We thank GenProbio srl for financial support of the Laboratory of Probiogenomics. LM is supported by Fondazione Cariparma, Parma, Italy. DvS is a member of The APC Microbiome Institute funded by Science Foundation Ireland (SFI), through the Irish Government's National Development Plan (Grant number SFI/12/RC/2273). Part of this research is conducted using the High Performance Computing (HPC) facility of the University of Parma.

Author Contributions

M.V. and T.M. conceived the study. A.T. and A.N. collected the samples and medical interpretation of the results. F.T., M.M. and A.V. performed libraries preparation and illumina sequencing. L.M., C.M., G.A.L. and D.v.S. performed the bioinformatics analyses and collaborated for the microbiological interpretation of the results. L.M., C.M., A.T., M.V., T.M. and D.v.S. wrote the paper.

Additional Information

Supplementary information accompanies this paper at doi:10.1038/s41598-017-10663-w

Competing Interests: The authors declare that they have no competing interests.

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

Chapter 7

Identification of universal gut microbial biomarkers of common human intestinal diseases by meta-analysis

Mancabelli L, Milani C, Lugli GA, Turroni F, Cocconi D, van Sinderen D and Ventura M

This chapter is under revision in FEMS Microbiology Ecology.

Identification of universal gut microbial biomarkers of common human intestinal diseases by meta-analysis.

Leonardo Mancabelli[1], Christian Milani[1], Gabriele Andrea Lugli[1], Francesca Turroni[1], Deborah Cocconi[1], Douwe van Sinderen[2] and Marco Ventura[1]

Laboratory of Probiogenomics, Department of Chemistry, Life Sciences and Environmental Sustainability, University of Parma, Parma, Italy[1]; APC Microbiome Institute and School of Microbiology, Bioscience Institute, National University of Ireland, Cork, Ireland[3]

1

Abstract

Intestinal diseases, such as Crohn's disease (CD), ulcerative colitis (UC) and pseudomembranous colitis (CDI), are among the most common diseases in humans and may lead to more serious pathologies, e.g. colorectal cancer (CRC). Next Generation Sequencing has in recent years allowed the identification of correlations between intestinal bacteria and diseases, although formulation of universal gut microbial biomarkers for such diseases is only in its infancy. In the current study, we selected and reanalyzed a total of 3048 public datasets obtained from 16S rRNA profiling of individuals affected by CD, UC, CDI and CRC. This meta-analysis revealed possible biases in the reconstruction of the gut microbiota composition due to the use of different primer pairs employed for PCR of 16S rRNA gene fragments. Notably, this approach also identified common features of individuals affected by gut diseases (DS), including lower biodiversity compared to control subjects (CTRL). Moreover, potential universal intestinal disease microbial biomarkers were identified through cross-disease comparisons. In detail, CTRL showed high abundance of the genera *Barnesiella*, *Ruminococcaceae* UCG-005, *Alistipes*, *Christensenellaceae* R-7 group and unclassified member of *Lachnospiraceae* family, while DS exhibited high abundance of Lactobacillus, unclassified member of *Erysipelotrichaceae* family and Streptococcus genera.

Introduction

In the last 50 years, the incidence and prevalence of inflammatory bowel diseases (IBD), such as Crohn's disease (CD) and ulcerative colitis (UC), has increased worldwide (Cosnes*, et al.*, 2011, Molodecky*, et al.*, 2012, Ananthakrishnan, 2015), especially in traditionally low-incident regions such as Asia, South America as well as southern and eastern Europe (Lovasz*, et al.*, 2013, Ng*, et al.*, 2013). Moreover, IBD represents one of the main risk factors for the development of colorectal cancer (CRC) (Kulaylat & Dayton, 2010) which is a major cause of morbidity and mortality throughout the world (Haggar & Boushey, 2009). Furthermore, populations living in industrialized countries have been affected in the last fifteen years by an increased incidence

of *Clostridium difficile* infections (CDI) (Reveles, *et al.*, 2014), representing one of the most common hospital-acquired infections, and being generally associated with antibiotic use and responsible for pseudomembranous colitis (Antharam, *et al.*, 2013). The recent development of high throughput sequencing technologies, such as Roche 454, Ion Torrent and Illumina, allowed profiling of the bacterial population harbored by the intestinal tract, i.e. gut microbiota, and characterization of alterations in this microbiota, sometimes referred to as gut dysbiosis, associated with major intestinal diseases (Carding, *et al.*, 2015).

Although many studies have investigated possible correlations between gut microbiota and intestinal diseases, thereby revealing possible gut bacterial biomarkers (Table S1), most of such studies would have focused on a single pathology such as CD (Perez-Brocal, *et al.*, 2015, Eun, *et al.*, 2016), UC (Duranti, *et al.*, 2016, Mar, *et al.*, 2016), CRC (Kostic, *et al.*, 2012, Geng, *et al.*, 2013, Wu, *et al.*, 2013, Zackular, *et al.*, 2014, Burns, *et al.*, 2015) and CDI (Yatsunenko, *et al.*, 2012, Antharam, *et al.*, 2013, Rojo, *et al.*, 2015, Gu, *et al.*, 2016, Khanna, *et al.*, 2016, Milani, *et al.*, 2016). Nonetheless, the accuracy of the majority of published microbiota analyses is very much dependent on the applied methodology (Turroni, *et al.*, 2012, Milani, *et al.*, 2013). One of the most critical steps for accurate 16S rRNA-based microbiota profiling is the selection of primer pairs used for amplification which may lead to under-representation or selection against single bacterial species or even complete microbial groups (Klindworth, *et al.*, 2013).

Here, we evaluate the accuracy of different and currently used 16S rRNA gene-targeting PCR primers, and evaluate their impact on the profiling of the gut microbiota. Furthermore, we performed a meta-analysis of case-control studies focusing on 16S rRNA profiling of the gut microbiota of individuals affected by CD, UC, CRC and CDI. We selected a total of 3048 datasets, 1252 corresponding to control subjects (CTRL) and 1796 corresponding to individuals with intestinal diseases, retrieved from 24 public studies (Table S2). In detail, we collected 359, 1457, 512 and 720 samples belonging to colorectal cancer (Kostic, *et al.*, 2012, Geng, *et al.*, 2013, Weir, *et al.*, 2013, Wu, *et al.*, 2013, Zackular, *et al.*, 2014, Burns, *et al.*, 2015), , Crohn's disease (Gevers, *et al.*, 2014, Perez-Brocal, *et al.*, 2015, Eun, *et al.*, 2016), Clostridium difficile infections (Yatsunenko, *et al.*, 2012, Antharam, *et al.*, 2013, Rojo, *et al.*, 2015, Gu, *et al.*, 2016, Khanna, *et al.*, 2016, Milani, *et al.*, 2016) and ulcerative colitis studies (Gevers, *et al.*, 2014, Duranti, *et al.*, 2016, Mar, *et al.*, 2016, Shah, *et al.*, 2016), respectively.

Materials and methods

Selection of databases. All datasets included in this meta-analysis were collected from publicly available and published comparative human gut microbiota studies in the context of Crohn's disease, ulcerative colitis, colorectal cancer and *Clostridium difficile* infections. For each intestinal disease, we collected 16S rRNA profiling datasets from a minimum of five studies. Illumina sequencing technology was preferred in order to ensure high data coverage and quality. Nevertheless, if Illumina datasets were not available, we included data produced by means of 454 sequencing. Moreover, selected datasets had to represent both control and diseased subjects, and obtained from fecal samples or biopsies collected from the adult human large intestine (average age of 22 ± 18).

Evaluation of primer pairs efficiency. The performance of primer pairs employed in the studies included in our meta-analysis (Table S3) were evaluated through the web-tool TestPrime 1.0 (Klindworth, *et al.*, 2013). The latter performs an *in silico* PCR using the SILVA database as template and provides the percentage of amplified sequences for each bacterial genus (Klindworth, *et al.*, 2013).

16S rRNA-based Microbiota Analysis. To avoid biases caused by different bioinformatic analysis pipelines, the sequence read pools of each study were filtered and analyzed using the same custom script based on the QIIME software suite (Caporaso, *et al.*, 2010). Quality control retained sequences with a length

between 140 and 400 bp and mean sequence quality score > 20 while sequences with homopolymers > 7 bp and mismatched primers were omitted. 16S rRNA Operational Taxonomic Units (OTUs) were defined at ≥ 97 % sequence homology using UCLUST (Edgar, 2010) and OTUs with less than 10 sequences were filtered. All reads were classified to the lowest possible taxonomic rank using QIIME (Caporaso, *et al.*, 2010) and a reference dataset from the SILVA database v.123 (Quast, *et al.*, 2013). In order to assess the bacterial complexity, the alpha-diversity was evaluated based on Chao1 indexes and represented by rarefaction curves generated using 10 subsampling of the whole datasets. Furthermore, the Bray-Curtis dissimilarity index was used to estimate the beta-diversity between CTRL and individuals affected by intestinal diseases. Dissimilarities were reported through a 3-Dimensional Principal Coordinate Analysis (PCoA) representation. QIIME and SPSS software (www.ibm.com/software/it/analytics/spss/) were used to compute statistical analyses. PERMANOVA were performed using 1,000 permutations to estimate p-values for differences among populations in PCoA analyses. Furthermore, differential abundance of bacterial genera and alpha-diversity was tested by ANOVA. Moreover, covariance analysis between primer pairs and bacterial relative abundance was performed through Pearson correlation coefficient.

Results and discussion

Homogeneity of the samples. 16S rRNA-based microbiota profiling is a technique that relies on next-generation sequencing data for a cost-effective analysis of the bacterial community present in a given environmental sample. Due to its accuracy and ability to profile non-cultivable taxa, 16S rRNA-based profiling rapidly became the most widely exploited approach for gut microbiota characterization. Nevertheless, absence of a gold standard protocol led to extensive methodological variation, with consequent output biases that might prevent reliable and meaningful comparisons between datasets derived from different studies (Milani, *et al.*, 2013). In detail, possible biases may be due to study design, sample collection, transport and storage of the samples, DNA extraction and other variables related to sequencing and bioinformatics analyses (Milani, *et al.*, 2013). Among the main reasons for variable data outputs is the species-specific efficiency and accuracy of the various sets of PCR primers employed to amplify part of the 16S rRNA genes that represent a given sample community (Milani, *et al.*, 2013). In order to evaluate the accuracy and efficacy of the 12 primer pairs that are used in the selected datasets, and that also represent the currently most frequently used PCR primers in 16S microbial profiling, we tested the primer pairs through *in silico* PCR. Notably, this assessment revealed rather variable amplification performances that are expected to cause genera-specific biases (Table S4).

***In silico* evaluation of the PCR primers accuracy**. Primer pairs Probio_Uni/Probio_Rev, 357F/926R, 338F/806R, 530F/926R, V5F/V6R, 341F/534R and 515F/806R showed an *in silico* efficacy of > 90 % in their ability to amplify the targeted 16S rRNA gene sequences. In contrast, primer pairs 27F/338R, 8F/357R, 8F/518R, 27F/534R and 8F/530R exhibited a predicted capacity of < 32 % for their ability to amplify their specifically targeted 16S rRNA sequences. We then focused on the evaluation of genus-specific amplification performances of intestinal taxa that had been determined to be present at a relative abundance of > 2% in at least one sample included in our meta-analysis (Table S5). Analysis of the 252 selected intestinal genera confirmed the efficiency observed regarding all bacterial 16S rRNA sequences. In detail, the bacterial sequences belonging to 248 genera were amplified by all primers examined and the primer pairs used in studies based on 454 sequencing showed the lowest efficiency (< 32 %) with the exception of 341F/534R (Muyzer, *et al.*, 1993, Juck, *et al.*, 2000) (efficiency at genus level of 95.54 %), 357F/926R (Liu, *et al.*, 1997) (efficiency at genus level of 95.38 %) and 515F/806R (Caporaso, *et al.*, 2011, Walters, *et al.*, 2011) (efficiency at genus level

of 95.24 %). Moreover, evaluation of the primer pair-mediated amplication efficiency for each of the bacterial sequences harboring the 252 selected taxa showed that 27F/338R (Hongoh, *et al.*, 2003, Fierer, *et al.*, 2008), 8F/530R (Frank, *et al.*, 2007, Perez-Brocal, *et al.*, 2013), 8F/357R, 27F/534R (Ben-Dov, *et al.*, 2006) and 8F/518R (Muyzer, *et al.*, 1993, Frank, *et al.*, 2007) primer pairs elicited an amplification efficacy of > 70 % only for the genera *Ethanoligenens*, *Fretibacterium* and *Lachnospiraceae* UCG-008. Notably, Probio_Uni/Probio_Rev (Milani, *et al.*, 2013) showed the highest *in silico* predicted PCR performances amongst all evaluated PCR primer pairs. In fact, the Probio_Uni/Probio_Rev primer pair was predicted to amplify the 16S rRNA gene sequences of 75.40 % of the 252 selected genera with an efficiency of > 95 %, followed by primer pairs 341F/534R (75.00 %) and 357F/926R (71.03 %). Furthermore, 530F/926R (Liu, *et al.*, 1997, Dowd, *et al.*, 2008), 515F/806R (Caporaso, *et al.*, 2011, Walters, *et al.*, 2011), V5F/V6R (Cai, *et al.*, 2013) and 338F/806R (el Fantroussi, *et al.*, 1999, Walters, *et al.*, 2011) displayed an efficiency > 95 % in less than 70 % of the assessed 252 genera. In order to evaluate the correlation between a given primer pair and corresponding predicted relative abundance at genus level, we performed a covariance analysis through Pearson correlation coefficient based on the 3048 datasets and primer pair efficiency. This analysis indicated that 50 genera displayed a positive correlation to a given primer pair-mediated amplification efficiency (p-value < 0.05), thereby indicating that the primer pair in question plays an important role in the generation of a bias in the determination of gut microbiota composition. Notably, when focusing on taxa with a relative abundance of > 0.1% at least one dataset (Figure 1), the primer pairs appear to have an impact on assessing the presence and abundance of certain taxa that are considered key gut commensal bacteria, such as *Bifidobacterium*, *Coprococcus 3*, genera belonging to *Ruminococcaceae* and *Eubacteriaceae*. These marked differences in amplification performance obtained for the tested 12 primer pairs therefore highlight the existence of biases in the reconstruction of the gut microbiota composition as reported by many published studies (Perez-Brocal, *et al.*, 2015, Rojo, *et al.*, 2015). This finding unfortunately prevents a reliable cross-study meta-analysis of all datasets corresponding to case and control subjects produced by different research projects. For this reason, each case-control sample processed with different laboratory protocols from several intestinal diseases, i.e. CD, UC, CRC and CDI, were analyzed separately. Subsequently, the study-specific results were evaluated together to define a global trend (increase or decrease) for each bacterial taxon in control versus disease condition.

Meta- and cross-analysis of the gut microbiota in intestinal diseases. Quality-filtering of CD, UC, CRC and CDI samples produced an average of 49,651, 66,127, 62,242 and 376,768 reads, respectively (Table S2). This level of DNA sequencing depth is considered appropriate to infer a thorough analysis of the gut microbiota (Hamady & Knight, 2009).

Analysis of the microbiota complexity evaluated through alpha-diversity cross-study meta-averages, e.g. averages of all the CTRL and all affected subjects for each intestinal disease analyzed, showed higher complexity in control samples compared to CD (p-value < 0.01) (Figure 2a), CRC (p-value < 0.05) (Figure 4a) and CDI (p-value < 0.01) (Figure 5a) samples. In addition, the meta-analyzed studies of UC samples provided very different alpha-diversity curves and, as expected, evaluation of the control and UC cross-study meta-averages showed a p-value > 0.05. Therefore, such data may indicate that biases in taxonomic reconstruction induced by the use of different analytical protocols, such as selection of primer pairs, significantly impact on the observed biodiversity (Hamady & Knight, 2009) thus precluding cross-study meta-analysis of alpha-diversity (Figure 3a).

Moreover, cross-study meta-PERMANOVA, i.e. PERMANOVA obtained for all CTRL samples and all affected subjects for each intestinal disease of the meta-analyzed studies,

based on the Bray-Curtis dissimilarity index showed a p-value < 0.001 for all comparisons indicating a taxonomical difference among the samples from control and diseased subjects (Figures 2b, 3b, 4c and 5b).

Cross-study meta-ANOVA of the bacterial profile at phylum level, e.g. averages of all CTRL and all affected subjects for each intestinal disease analyzed, showed predominant abundance of *Bacteroidetes* in control samples in mainly disease analyzed, i.e. average of 44.42 % (p-value < 0.01), 45.43 % (p-value < 0.01) and 34.01 % (p-value < 0.01) compared to CD, UC and CRC, respectively (Figures 2c, 3c and 4c). Conversely, when compared to control samples the gut microbiota of diseased samples appears to exhibit a higher abundance of the *Proteobacteria* phylum, such as in the case of samples from individuals suffering from CD (average 19.11 %, p-value < 0.01), CRC (average 15.11 %, p-value < 0.01) and CDI (average 22.49 %, p-value < 0.01) (Figures 2c, 4c and 5c).

Comparison between the gut microbiota composition of control individuals and of subjects for each intestinal disease analyzed at genus level, showed higher abundance of genera belonging to the *Bacteroidetes* phylum in CTRL samples (Figure 2d, 3d, 4d, 5d). In contrast, genera belonging to *Proteobacteria* were particularly less abundant in control samples as compared to samples corresponding to each of the investigated diseases (Figures 2d, 3d, 4d, 5d).

Interestingly, comparison between the gut microbiota composition of CD and CTRL samples showed higher abundance of genera *Parabacteroides* (increase of 67.10 %, p-value < 0.05), *Faecalibacterium* (increase of 18.19 %, p-value < 0.05), *Prevotella 9* (increase of 185.24 %, p-value < 0.05) and *Bacteroides* (increase of 45.68 %, p-value < 0.05), in CTRL samples (Figure 2d). In contrast, genera *Escherichia-Shigella* (decrease of -39.08 %, p-value < 0.05) and *Haemophilus* (decrease of -39.08 %, p-value < 0.05) were less abundant in control samples as compared to samples obtained from individuals with CD (Figure 2d). Notably, CRC samples possess a higher abundance of bacteria that have been associated with the development of intestinal diseases, such as *Campylobacter* (increase of 950.04 %, p-value < 0.05) (Warren*, et al.*, 2013, Akutko & Matusiewicz, 2017), or known to be involved in the transition from eubiosis, i.e. an optimal balance of gut microbiota composition (Iebba*, et al.*, 2016), to dysbiosis, such as *Gemella* (increase of 118.05 %, p-value < 0.05) (Chen*, et al.*, 2016) (Figure 4d). Interestingly, 16S microbial profiles of CTRL samples displayed a higher abundance of members of the genus *Faecalibacterium* (increase of 77.91 %, p-value < 0.05), which is considered a bacterial genus with a beneficial effect on the human gut (Ventura*, et al.*, 2014) and which could have a role in preventing CRC (Wei*, et al.*, 2016). Interestingly, also the gut microbiota profiles of CDI samples possess a higher abundance of opportunistic pathogens belonging to the phylum *Proteobacteria* and a lower abundance of taxa that are associated with health promoting effects, such as *Bifidobacterium* and *Faecalibacterium* (Milani*, et al.*, 2014, Ventura*, et al.*, 2014, Milani*, et al.*, 2015) (Figure 5d).

Moreover, evaluation of the taxonomic trend (Tables S7, S9, S11 and S13) and the differences of the gut microbiota composition at genus level across the meta-analyzed studies allowed us to identify genera that may represent suitable bacterial biomarkers of each analyzed disease. Interestingly, the relative abundance of 16, 5, 7 and 3 taxa increase wile 2, 4, 4 and 3 taxa decrease respectively in CD, UC, CRC and CDI subjects when compared to CTRL individuals in in all meta-analyzed studies (Table S7, Table S9, Table S11 and Table S13). A summary of the taxa that may constitute specific disease microbial biomarkers were reported in Table 1.

Identification of universal biomarkers. In order to evaluate the existence of universal intestinal diseases biomarkers, we performed a meta-analysis for all datasets corresponding to studies regarding CD, UC, CRC and CDI. Cross-analysis of the alpha-diversity showed a higher biodiversity in CTRL samples with respect to subjects affected by an intestinal disease (DS) (p-value < 0.05) (Figure 6a). These data confirm previous observations that

intestinal dysbiosis is linked to loss of microbiota diversity (Sha, *et al.*, 2013, Mosca, *et al.*, 2016). Moreover, the beta-diversity cross-analysis indicated a clear division between CTRL and samples affected by intestinal diseases (meta-PERMANOVA p-value < 0.05) (Figure 6b). Therefore, we focused at the genus level to identify the differences in gut microbiota composition between these two groups. In detail, these analyses revealed a total of 261 genera with a significantly different abundance (p-value < 0.05) (Table S14), of which 20 with an average abundance variation of > 0.5 % (Figure 6c).

Interestingly, when focusing on the genera with a significant p-value and a taxonomic trend with a prevalence of > 80 % (Table S15), it was possible to identify five and three taxa characteristic of CTRL and DS subjects, respectively. In detail, CTRL showed high relative abundance of the genera *Barnesiella* (in 90 % of the studies and p-value < 0.05), *Ruminococcaceae* UCG-005 (in 85 % of the studies and p-value < 0.05), *Alistipes* (in 80 % of the studies and p-value < 0.05), *Christensenellaceae* R-7 group (in 80 % of the studies and p-value < 0.05) and unclassified member of *Lachnospiraceae* family (in 80 % of the studies and p-value < 0.05), while DS displayed high abundance of the taxa *Lactobacillus* (in 90 % of the studies and p-value < 0.05), unclassified member of *Erysipelotrichaceae* family (in 80 % of the studies and p-value < 0.05) and *Streptococcus* (in 80 % of the studies and p-value < 0.05).

In previous studies, *Barnesiella* genus was identified only in populations living in developed countries (Mancabelli, *et al.*, 2017) and was correlated to beneficial effects on human gut (Kulagina, *et al.*, 2012, Ubeda, *et al.*, 2013). Moreover, *Ruminococcaceae* UCG-005, *Alistipes*, and unclassified member of *Lachnospiraceae* family have been reported to be butyrate-producing bacteria (Flint, *et al.*, 2012, Chen, *et al.*, 2017) that may protect healthy subjects from chronic intestinal inflammation (Lepage, *et al.*, 2011). In contrast, the higher relative abundance of *Streptococcus* genus in DS confirm its previously-reported correlation to a range of gastrointestinal diseases (Murray & Roberts, 1978, Burnett-Hartman, *et al.*, 2008) and renders it a valuable candidate as a universal biomarker of intestinal dysbiosis. Furthermore, bacteria belonging to *Erysipelotrichaceae* family were correlated to inflammation (Kaakoush, 2015) and immunomodulation (Palm, *et al.*, 2014) but their functional correlation to intestinal diseases is far from being fully elucidated.

Notably, the observed higher relative abundance of the non-pathogenic taxa *Lactobacillus* in DS may reflect lower niche-competition caused by simplification of the dysbiotic gut microbiota (Walter, 2008).

CONCLUSIONS

A substantial number of studies based on 16S rRNA gene profiling have reported on the correlation between human gut diseases and microbiota composition. Nevertheless, one of the main biases in the reconstruction of the gut microbiota composition through 16S rRNA profiling is the selection of reliable and universal primer pairs. *In silico* PCR and covariance analysis of the 12 primer pairs used in 24 selected public gut metagenomic studies confirmed their impact on biased amplification of the targeted section of the 16S rRNA gene. To overcome this limitation, we performed a cross-study meta-analysis of 3048 public metagenomic datasets, corresponding to 1252 control (CTRL) and 1796 patient subjects, in order to identify possible bacterial biomarkers for major intestinal diseases such as CD, UC, CRC and CDI. Furthermore, we analyzed all datasets together, in order to identify possible universal gut disease microbial biomarkers. In detail, this cross-study analysis showed that *Barnesiella*, *Ruminococcaceae* UCG-005, *Alistipes*, *Christensenellaceae R-7* group and unclassified member of *Lachnospiraceae* family genera correlated with a healthy state of subjects. In contrast, subjects that present an intestinal disease displayed higher abundance of genera reported to cause intestinal inflammation, such as unclassified member of *Erysipelotrichaceae* family and *Streptococcus*. The identification of novel universal biomarkers as indicators of human gut diseases may contribute to rapid diagnosis as well as to predict the course and prognosis of the disease

and guide therapeutic decisions improving patient care.

Acknowledgements
Part of this research is conducted using the High Performance Computing (HPC) facility of the University of Parma. This work was funded by the EU Joint Programming Initiative – A Healthy Diet for a Healthy Life (JPI HDHL, http://www.healthydietforhealthylife.eu/) to MV and DvS (Grant no. 15/JP/HDHL/3280), and the MIUR to MV. We thank GenProbio srl for financial support of the Laboratory of Probiogenomics. LM is supported by Fondazione Cariparma, Parma, Italy. DvS is a member of The APC Microbiome Institute funded by Science Foundation Ireland (SFI), through the Irish Government's National Development Plan (Grant no. SFI/12/RC/2273). The authors declare that no competing interests exist.

References

Akutko K & Matusiewicz K (2017) Campylobacter concisus as the etiologic agent of gastrointestinal diseases. *Adv Clin Exp Med* **26**: 149-154.

Ananthakrishnan AN (2015) Epidemiology and risk factors for IBD. *Nat Rev Gastroenterol Hepatol* **12**: 205-217.

Antharam VC, Li EC, Ishmael A, Sharma A, Mai V, Rand KH & Wang GP (2013) Intestinal dysbiosis and depletion of butyrogenic bacteria in Clostridium difficile infection and nosocomial diarrhea. *J Clin Microbiol* **51**: 2884-2892.

Ben-Dov E, Shapiro OH, Siboni N & Kushmaro A (2006) Advantage of using inosine at the 3' termini of 16S rRNA gene universal primers for the study of microbial diversity. *Appl Environ Microbiol* **72**: 6902-6906.

Burnett-Hartman AN, Newcomb PA & Potter JD (2008) Infectious agents and colorectal cancer: a review of Helicobacter pylori, Streptococcus bovis, JC virus, and human papillomavirus. *Cancer Epidemiol Biomarkers Prev* **17**: 2970-2979.

Burns MB, Lynch J, Starr TK, Knights D & Blekhman R (2015) Virulence genes are a signature of the microbiome in the colorectal tumor microenvironment. *Genome Med* **7**: 55.

Cai L, Ye L, Tong AH, Lok S & Zhang T (2013) Biased diversity metrics revealed by bacterial 16S pyrotags derived from different primer sets. *PLoS One* **8**: e53649.

Caporaso JG, Lauber CL, Walters WA*, et al.* (2011) Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc Natl Acad Sci U S A* **108 Suppl 1**: 4516-4522.

Caporaso JG, Kuczynski J, Stombaugh J*, et al.* (2010) QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* **7**: 335-336.

Carding S, Verbeke K, Vipond DT, Corfe BM & Owen LJ (2015) Dysbiosis of the gut microbiota in disease. *Microb Ecol Health Dis* **26**: 26191.

Chen T, Long W, Zhang C, Liu S, Zhao L & Hamaker BR (2017) Fiber-utilizing capacity varies in Prevotella- versus Bacteroides-dominated gut microbiota. *Sci Rep* **7**: 2594.

Chen Y, Ji F, Guo J, Shi D, Fang D & Li L (2016) Dysbiosis of small intestinal microbiota in liver cirrhosis and its association with etiology. *Sci Rep* **6**: 34055.

Cosnes J, Gower-Rousseau C, Seksik P & Cortot A (2011) Epidemiology and natural history of inflammatory bowel diseases. *Gastroenterology* **140**: 1785-1794.

Dowd SE, Callaway TR, Wolcott RD, Sun Y, McKeehan T, Hagevoort RG & Edrington TS (2008) Evaluation of the bacterial diversity in the feces of cattle using 16S rDNA bacterial tag-encoded FLX amplicon pyrosequencing (bTEFAP). *BMC Microbiol* **8**: 125.

Duranti S, Gaiani F, Mancabelli L*, et al.* (2016) Elucidating the gut microbiome of ulcerative colitis: bifidobacteria as novel microbial biomarkers. *FEMS Microbiol Ecol* **92**.

Edgar RC (2010) Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**: 2460-2461.

el Fantroussi S, Verschuere L, Verstraete W & Top EM (1999) Effect of phenylurea herbicides on soil microbial communities estimated by analysis of 16S rRNA gene fingerprints and community-level physiological profiles. *Appl Environ Microbiol* **65**: 982-988.

Eun CS, Kwak MJ, Han DS*, et al.* (2016) Does the intestinal microbial community of Korean Crohn's disease patients differ from that of western patients? *BMC Gastroenterol* **16**: 28.

Fierer N, Hamady M, Lauber CL & Knight R (2008) The influence of sex, handedness, and washing on the diversity of hand surface bacteria. *Proc Natl Acad Sci U S A* **105**: 17994-17999.

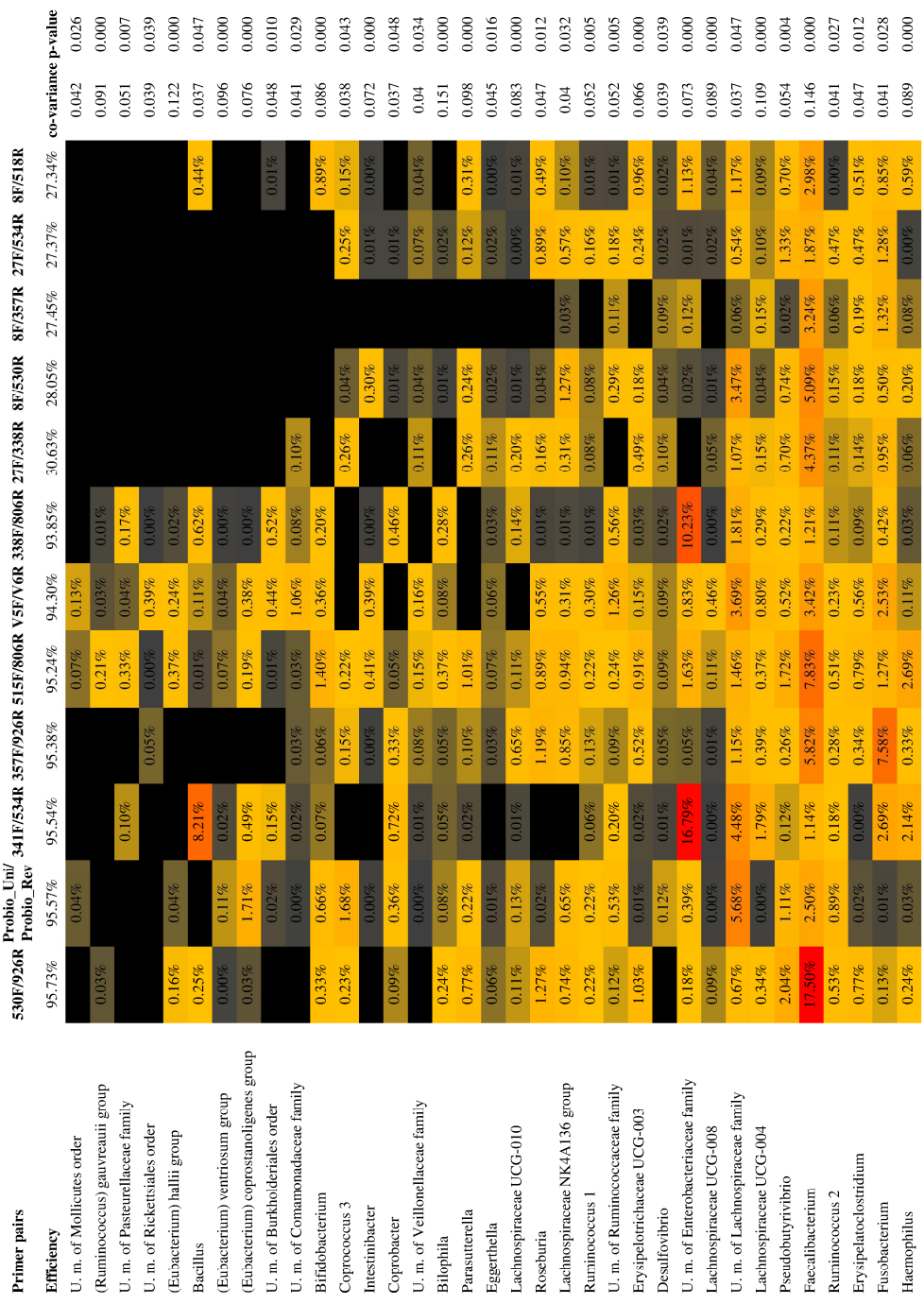
Flint HJ, Scott KP, Duncan SH, Louis P & Forano E (2012) Microbial degradation of complex carbohydrates in the gut. *Gut Microbes* **3**: 289-306.

Frank DN, St Amand AL, Feldman RA, Boedeker EC, Harpaz N & Pace NR (2007) Molecular-phylogenetic characterization of microbial community imbalances in human inflammatory bowel diseases. *Proc Natl Acad Sci U S A* **104**: 13780-13785.

Geng J, Fan H, Tang X, Zhai H & Zhang Z (2013) Diversified pattern of the human colorectal cancer microbiome. *Gut Pathog* **5**: 2.
Gevers D, Kugathasan S, Denson LA*, et al.* (2014) The treatment-naive microbiome in new-onset Crohn's disease. *Cell Host Microbe* **15**: 382-392.
Gu S, Chen Y, Zhang X*, et al.* (2016) Identification of key taxa that favor intestinal colonization of Clostridium difficile in an adult Chinese population. *Microbes Infect* **18**: 30-38.
Haggar FA & Boushey RP (2009) Colorectal cancer epidemiology: incidence, mortality, survival, and risk factors. *Clin Colon Rectal Surg* **22**: 191-197.
Hamady M & Knight R (2009) Microbial community profiling for human microbiome projects: Tools, techniques, and challenges. *Genome Res* **19**: 1141-1152.
Hongoh Y, Ohkuma M & Kudo T (2003) Molecular analysis of bacterial microbiota in the gut of the termite Reticulitermes speratus (Isoptera; Rhinotermitidae). *FEMS Microbiol Ecol* **44**: 231-242.
Iebba V, Totino V, Gagliardi A*, et al.* (2016) Eubiosis and dysbiosis: the two sides of the microbiota. *New Microbiol* **39**: 1-12.
Juck D, Charles T, Whyte LG & Greer CW (2000) Polyphasic microbial community analysis of petroleum hydrocarbon-contaminated soils from two northern Canadian communities. *FEMS Microbiol Ecol* **33**: 241-249.
Kaakoush NO (2015) Insights into the Role of Erysipelotrichaceae in the Human Host. *Front Cell Infect Microbiol* **5**: 84.
Khanna S, Montassier E, Schmidt B, Patel R, Knights D, Pardi DS & Kashyap P (2016) Gut microbiome predictors of treatment response and recurrence in primary Clostridium difficile infection. *Aliment Pharmacol Ther* **44**: 715-727.
Klindworth A, Pruesse E, Schweer T, Peplies J, Quast C, Horn M & Glockner FO (2013) Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res* **41**: e1.
Kostic AD, Gevers D, Pedamallu CS*, et al.* (2012) Genomic analysis identifies association of Fusobacterium with colorectal carcinoma. *Genome Res* **22**: 292-298.
Kulagina EV, Efimov BA, Maximov PY, Kafarskaia LI, Chaplin AV & Shkoporov AN (2012) Species composition of Bacteroidales order bacteria in the feces of healthy people of various ages. *Biosci Biotechnol Biochem* **76**: 169-171.
Kulaylat MN & Dayton MT (2010) Ulcerative colitis and cancer. *J Surg Oncol* **101**: 706-712.
Lepage P, Hasler R, Spehlmann ME*, et al.* (2011) Twin study indicates loss of interaction between microbiota and mucosa of patients with ulcerative colitis. *Gastroenterology* **141**: 227-236.
Liu WT, Marsh TL, Cheng H & Forney LJ (1997) Characterization of microbial diversity by determining terminal restriction fragment length polymorphisms of genes encoding 16S rRNA. *Appl Environ Microbiol* **63**: 4516-4522.
Lovasz BD, Golovics PA, Vegh Z & Lakatos PL (2013) New trends in inflammatory bowel disease epidemiology and disease course in Eastern Europe. *Dig Liver Dis* **45**: 269-276.
Mancabelli L, Milani C, Lugli GA, Turroni F, Ferrario C, van Sinderen D & Ventura M (2017) Meta-analysis of the human gut microbiome from urbanized and pre-agricultural populations. *Environ Microbiol* **19**: 1379-1390.
Mar JS, LaMere BJ, Lin DL, Levan S, Nazareth M, Mahadevan U & Lynch SV (2016) Disease Severity and Immune Activity Relate to Distinct Interkingdom Gut Microbiome States in Ethnically Distinct Ulcerative Colitis Patients. *MBio* **7**.
Milani C, Hevia A, Foroni E*, et al.* (2013) Assessing the fecal microbiota: an optimized ion torrent 16S rRNA gene-based analysis protocol. *PLoS One* **8**: e68739.
Milani C, Lugli GA, Duranti S*, et al.* (2014) Genomic encyclopedia of type strains of the genus Bifidobacterium. *Appl Environ Microbiol* **80**: 6290-6302.
Milani C, Ticinesi A, Gerritsen J*, et al.* (2016) Gut microbiota composition and Clostridium difficile infection in hospitalized elderly individuals: a metagenomic study. *Sci Rep* **6**: 25945.

Milani C, Lugli GA, Duranti S, *et al.* (2015) Bifidobacteria exhibit social behavior through carbohydrate resource sharing in the gut. *Sci Rep* **5**: 15782.
Molodecky NA, Soon IS, Rabi DM, *et al.* (2012) Increasing incidence and prevalence of the inflammatory bowel diseases with time, based on systematic review. *Gastroenterology* **142**: 46-54 e42; quiz e30.
Mosca A, Leclerc M & Hugot JP (2016) Gut Microbiota Diversity and Human Diseases: Should We Reintroduce Key Predators in Our Ecosystem? *Front Microbiol* **7**: 455.
Murray HW & Roberts RB (1978) Streptococcus bovis bacteremia and underlying gastrointestinal disease. *Arch Intern Med* **138**: 1097-1099.
Muyzer G, de Waal EC & Uitterlinden AG (1993) Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. *Appl Environ Microbiol* **59**: 695-700.
Ng SC, Tang W, Ching JY, *et al.* (2013) Incidence and phenotype of inflammatory bowel disease based on results from the Asia-pacific Crohn's and colitis epidemiology study. *Gastroenterology* **145**: 158-165 e152.
Palm NW, de Zoete MR, Cullen TW, *et al.* (2014) Immunoglobulin A coating identifies colitogenic bacteria in inflammatory bowel disease. *Cell* **158**: 1000-1010.
Perez-Brocal V, Garcia-Lopez R, Nos P, Beltran B, Moret I & Moya A (2015) Metagenomic Analysis of Crohn's Disease Patients Identifies Changes in the Virome and Microbiome Related to Disease Status and Therapy, and Detects Potential Interactions and Biomarkers. *Inflamm Bowel Dis* **21**: 2515-2532.
Perez-Brocal V, Garcia-Lopez R, Vazquez-Castellanos JF, Nos P, Beltran B, Latorre A & Moya A (2013) Study of the viral and microbial communities associated with Crohn's disease: a metagenomic approach. *Clin Transl Gastroenterol* **4**: e36.
Quast C, Pruesse E, Yilmaz P, *et al.* (2013) The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* **41**: D590-596.
Reveles KR, Lee GC, Boyd NK & Frei CR (2014) The rise in Clostridium difficile infection incidence among hospitalized adults in the United States: 2001-2010. *Am J Infect Control* **42**: 1028-1032.
Rojo D, Gosalbes MJ, Ferrari R, *et al.* (2015) Clostridium difficile heterogeneously impacts intestinal community architecture but drives stable metabolome responses. *ISME J* **9**: 2206-2220.
Sha S, Xu B, Wang X, *et al.* (2013) The biodiversity and composition of the dominant fecal microbiota in patients with inflammatory bowel disease. *Diagn Microbiol Infect Dis* **75**: 245-251.
Shah R, Cope JL, Nagy-Szakal D, Dowd S, Versalovic J, Hollister EB & Kellermayer R (2016) Composition and function of the pediatric colonic mucosal microbiome in untreated patients with ulcerative colitis. *Gut Microbes* **7**: 384-396.
Turroni F, Peano C, Pass DA, *et al.* (2012) Diversity of bifidobacteria within the infant gut microbiota. *PLoS One* **7**: e36957.
Ubeda C, Bucci V, Caballero S, *et al.* (2013) Intestinal microbiota containing Barnesiella species cures vancomycin-resistant Enterococcus faecium colonization. *Infect Immun* **81**: 965-973.
Ventura M, Turroni F, Lugli GA & van Sinderen D (2014) Bifidobacteria and humans: our special friends, from ecological to genomics perspectives. *J Sci Food Agric* **94**: 163-168.
Walter J (2008) Ecological role of lactobacilli in the gastrointestinal tract: implications for fundamental and biomedical research. *Appl Environ Microbiol* **74**: 4985-4996.
Walters WA, Caporaso JG, Lauber CL, Berg-Lyons D, Fierer N & Knight R (2011) PrimerProspector: de novo design and taxonomic analysis of barcoded polymerase chain reaction primers. *Bioinformatics* **27**: 1159-1161.
Warren RL, Freeman DJ, Pleasance S, *et al.* (2013) Co-occurrence of anaerobic bacteria in colorectal carcinomas. *Microbiome* **1**: 16.
Wei Z, Cao S, Liu S, *et al.* (2016) Could gut microbiota serve as prognostic biomarker associated with colorectal cancer patients' survival? A pilot study on relevant mechanism. *Oncotarget* **7**: 46158-46172.

Weir TL, Manter DK, Sheflin AM, Barnett BA, Heuberger AL & Ryan EP (2013) Stool microbiome and metabolome differences between colorectal cancer patients and healthy adults. *PLoS One* **8**: e70803.
Wu N, Yang X, Zhang R*, et al.* (2013) Dysbiosis signature of fecal microbiota in colorectal cancer patients. *Microb Ecol* **66**: 462-470.
Yatsunenko T, Rey FE, Manary MJ*, et al.* (2012) Human gut microbiome viewed across age and geography. *Nature* **486**: 222-227.
Zackular JP, Rogers MA, Ruffin MTt & Schloss PD (2014) The human gut microbiome as a screening tool for colorectal cancer. *Cancer Prev Res (Phila)* **7**: 1112-1121.



Primer pairs	530F/926R	Probio_Uni/ Probio_Rev	341F/534R	357F/926R	515F/806R	V5F/V6R	338F/806R	27F/338R	8F/530R	8F/357R	27F/534R	8F/518R	co-variance	p-value
Efficiency	95.73%	95.57%	95.54%	95.38%	95.24%	94.30%	93.85%	30.63%	28.05%	27.45%	27.37%	27.34%		
U. m. of Mollicutes order		0.04%			0.07%	0.13%							0.042	0.026
(Ruminococcus) gauvreauii group	0.03%				0.21%	0.03%	0.01%						0.091	0.000
U. m. of Pasteurellaceae family			0.10%		0.33%	0.04%	0.17%						0.051	0.007
U. m. of Rickettsiales order				0.05%	0.00%	0.39%	0.00%						0.039	0.039
(Eu)acterium) hallii group	0.16%	0.04%			0.37%	0.24%	0.02%						0.122	0.000
Bacillus	0.25%		8.21%		0.01%	0.11%	0.62%					0.44%	0.037	0.047
(Eu)acterium) ventriosum group	0.00%	0.11%	0.02%		0.07%	0.04%	0.00%						0.096	0.000
(Eu)acterium) coprostanoligenes group	0.03%	1.71%	0.49%		0.19%	0.38%	0.00%						0.076	0.000
U. m. of Burkholderiales order		0.02%	0.15%		0.01%	0.44%	0.52%					0.01%	0.048	0.010
U. m. of Comamonadaceae family		0.00%	0.02%	0.03%	0.03%	1.06%	0.08%	0.10%					0.041	0.029
Bifidobacterium	0.33%	0.66%	0.07%	0.06%	1.40%	0.36%	0.20%					0.89%	0.086	0.000
Coprococcus 3	0.23%	1.68%		0.15%	0.22%			0.26%	0.04%		0.25%	0.15%	0.038	0.043
Intestinibacter		0.00%		0.00%	0.41%	0.39%	0.00%		0.30%		0.01%	0.00%	0.072	0.000
Coprobacter	0.09%	0.36%	0.72%	0.33%	0.05%		0.46%		0.01%		0.01%		0.037	0.048
U. m. of Veillonellaceae family		0.00%	0.01%	0.08%	0.15%	0.16%		0.11%	0.04%		0.07%	0.04%	0.04	0.034
Bilophila	0.24%	0.08%	0.05%	0.05%	0.37%	0.08%	0.28%		0.01%		0.02%		0.151	0.000
Parasutterella	0.77%	0.22%	0.02%	0.10%	1.01%			0.26%	0.24%		0.12%	0.31%	0.098	0.000
Eggerthella	0.06%	0.01%		0.03%	0.07%	0.06%	0.03%	0.11%	0.02%		0.02%	0.00%	0.045	0.016
Lachnospiraceae UCG-010	0.11%	0.13%	0.01%	0.65%	0.11%		0.14%	0.20%	0.01%		0.00%	0.01%	0.083	0.000
Roseburia	1.27%	0.02%		1.19%	0.89%	0.55%	0.01%	0.16%	0.04%		0.89%	0.49%	0.047	0.012
Lachnospiraceae NK4A136 group	0.74%	0.65%		0.85%	0.94%	0.31%	0.01%	0.31%	1.27%	0.03%	0.57%	0.10%	0.04	0.032
Ruminococcus 1	0.22%	0.22%	0.06%	0.13%	0.22%	0.30%	0.01%	0.08%	0.08%		0.16%	0.01%	0.052	0.005
U. m. of Ruminococcaceae family	0.12%	0.53%	0.20%	0.09%	0.24%	1.26%	0.56%		0.29%	0.11%	0.18%	0.01%	0.052	0.005
Erysipelotrichaceae UCG-003	1.03%	0.01%	0.02%	0.52%	0.91%	0.15%	0.03%	0.49%	0.18%		0.24%	0.96%	0.066	0.000
Desulfovibrio		0.12%	0.01%	0.05%	0.09%	0.09%	0.02%	0.10%	0.04%	0.09%	0.02%	0.02%	0.039	0.039
U. m. of Enterobacteriaceae family	0.18%	0.39%	16.79%	0.05%	1.63%	0.83%	10.23%		0.02%	0.12%	0.01%	1.13%	0.073	0.000
Lachnospiraceae UCG-008	0.09%	0.00%	0.00%	0.01%	0.11%	0.46%	0.00%	0.05%	0.01%		0.02%	0.04%	0.089	0.000
U. m. of Lachnospiraceae family	0.67%	5.68%	4.48%	1.15%	1.46%	3.69%	1.81%	1.07%	3.47%	0.06%	0.54%	1.17%	0.037	0.047
Lachnospiraceae UCG-004	0.34%	0.00%	1.79%	0.39%	0.37%	0.80%	0.29%	0.15%	0.04%	0.15%	0.10%	0.09%	0.109	0.000
Pseudobutyrivibrio	2.04%	1.11%	0.12%	0.26%	1.72%	0.52%	0.22%	0.70%	0.74%	0.02%	1.33%	0.70%	0.054	0.004
Faecalibacterium	17.50%	2.50%	1.14%	5.82%	7.83%	3.42%	1.21%	4.37%	5.09%	3.24%	1.87%	2.98%	0.146	0.000
Ruminococcus 2	0.53%	0.89%	0.18%	0.28%	0.51%	0.23%	0.11%	0.11%	0.15%	0.06%	0.47%	0.00%	0.041	0.027
Erysipelatoclostridium	0.77%	0.02%	0.00%	0.34%	0.79%	0.56%	0.09%	0.14%	0.18%	0.19%	0.47%	0.51%	0.047	0.012
Fusobacterium	0.13%	0.01%	2.69%	7.58%	1.27%	2.53%	0.42%	0.95%	0.50%	1.32%	1.28%	0.85%	0.041	0.028
Haemophilus	0.24%	0.03%	2.14%	0.33%	2.69%	0.11%	0.03%	0.06%	0.20%	0.08%	0.00%	0.59%	0.089	0.000

Figure 1. Covariance analysis based on the selected public datasets and primer pair efficiency. A heat map illustrating relative abundance of genera with a significant positive correlation to primer pair efficiency was shown. Only the genera with a relative abundance > 0.1% in at least one dataset were reported.

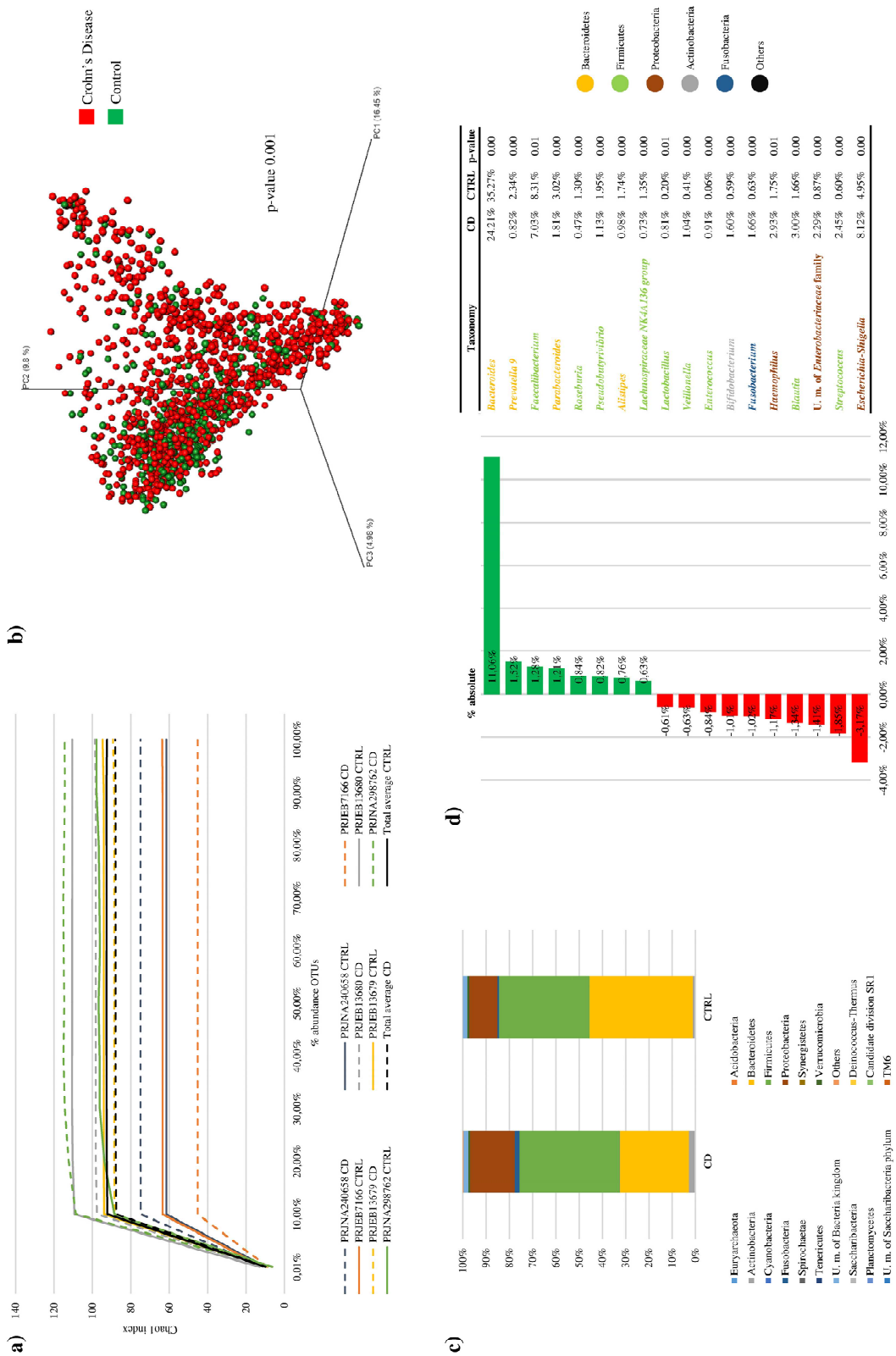


Figure 2. Exploration of the diversity and bacterial composition of CD and CTRL samples. Panel a shows the rarefaction curves calculated through Chao1 index of each study and of CD and CTRL average. Panel b reports the principal coordinate analysis (PCoA) of the collected case-control belonging to CD studies. Panel c displays the bacterial composition at phylum level based on cross-study of the CD and CTRL groups. Panel d reports the bacterial genera present an average abundance variation of > 0.5 %.

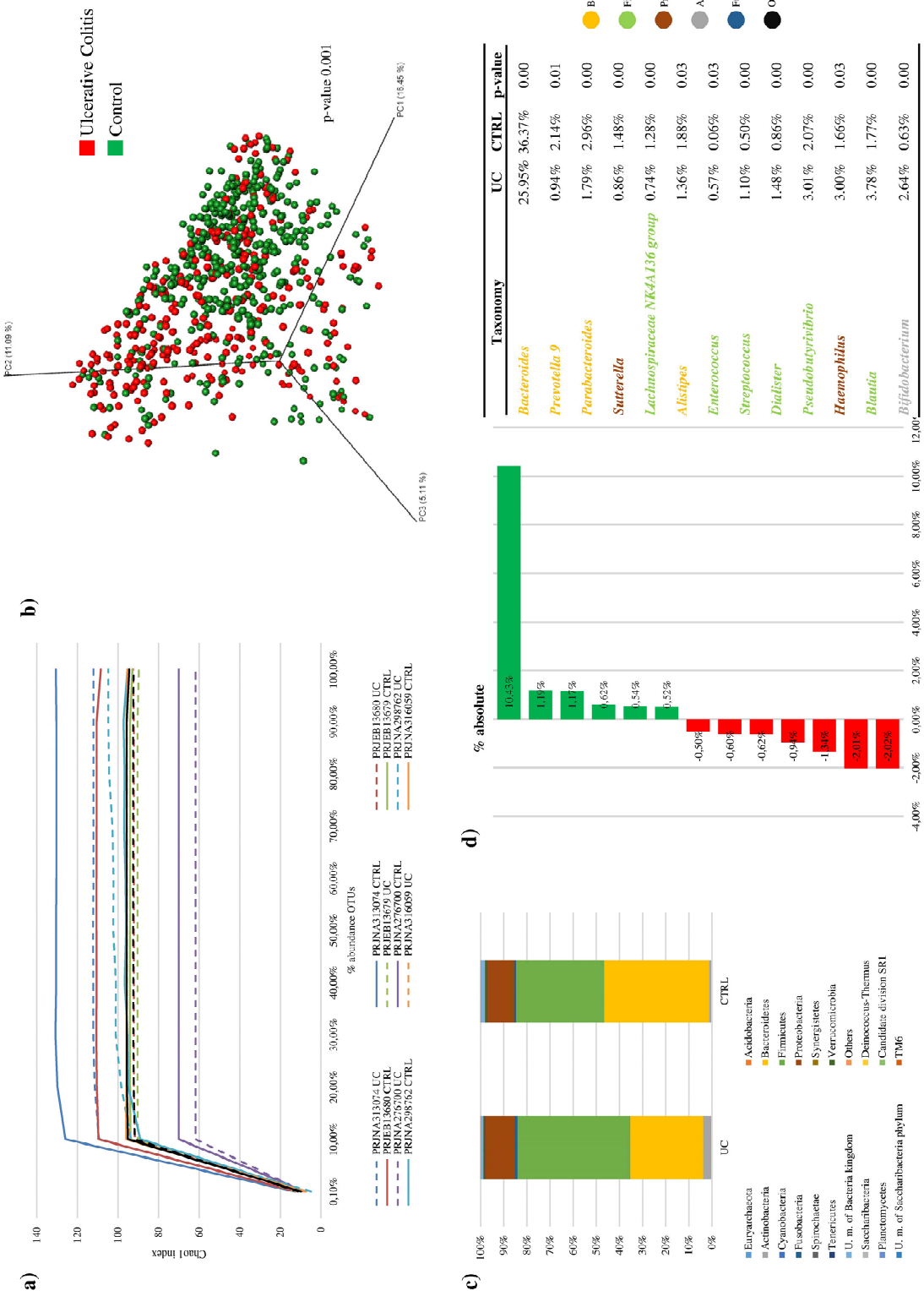
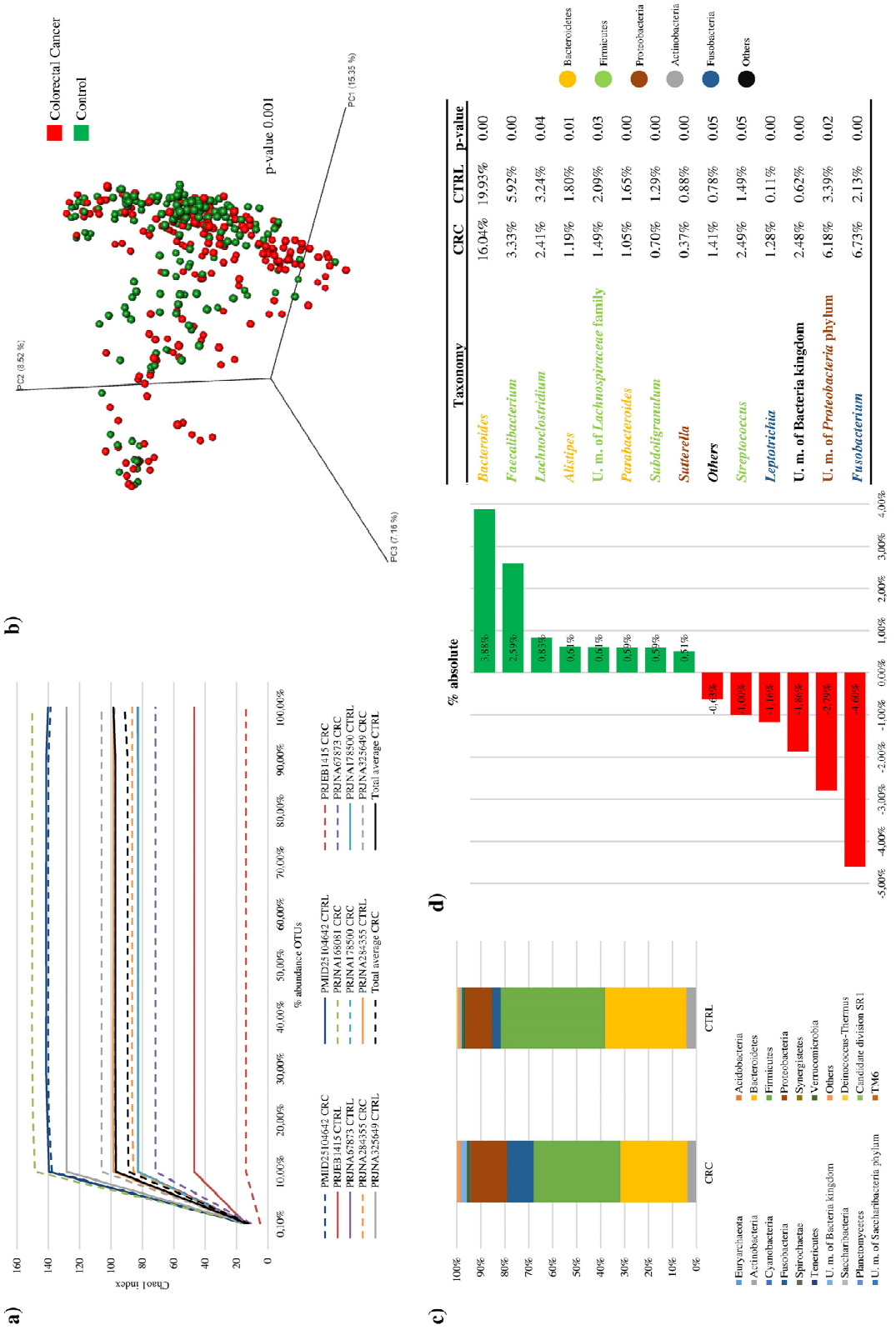


Figure 3. Evaluation of the microbiota composition of UC and CTRL samples. Panel a indicates the alpha-diversity curves calculated through Chao1 index of each study and of UC and CTRL average. Panel b shows the PCoA of the UC and CTRL groups. Panel c reports a bar plot depicting the bacterial composition at phylum level based on cross-study of the UC and CTRL groups. Panel d represents the average abundance variation > 0.5 % at genus level.



Taxonomy	CRC	CTRL	p-value	
Bacteroides	16.04%	19.93%	0.00	Bacteroidetes
Faecalibacterium	3.33%	5.92%	0.00	
Lachnoclostridium	2.41%	3.24%	0.04	
Alistipes	1.19%	1.80%	0.01	
U. m. of *Lachnospiraceae* family	1.49%	2.09%	0.03	Firmicutes
Parabacteroides	1.05%	1.65%	0.00	Proteobacteria
Subdoligranulum	0.70%	1.29%	0.00	
Sutterella	0.37%	0.88%	0.00	Actinobacteria
Others	1.41%	0.78%	0.05	Fusobacteria
Streptococcus	2.49%	1.49%	0.05	Others
Leptotrichia	1.28%	0.11%	0.00	
U. m. of Bacteria kingdom	2.48%	0.62%	0.00	
U. m. of *Proteobacteria* phylum	6.18%	3.39%	0.02	
Fusobacterium	6.73%	2.13%	0.00	

Figure 4. Examination of the complexity and bacterial differences between CRC and CTRL samples. Panel a displays the rarefaction curves of each study and of CRC and CTRL average. Panel b reports the beta-diversity of the collected case-control belonging to CRC studies. Panel c shows the bacterial composition at phylum level based on cross-study of the CRC and CTRL groups. Panel d reports the microbiota composition at genus level. The panel reports only the taxa that present an average abundance difference > 0.5 %.

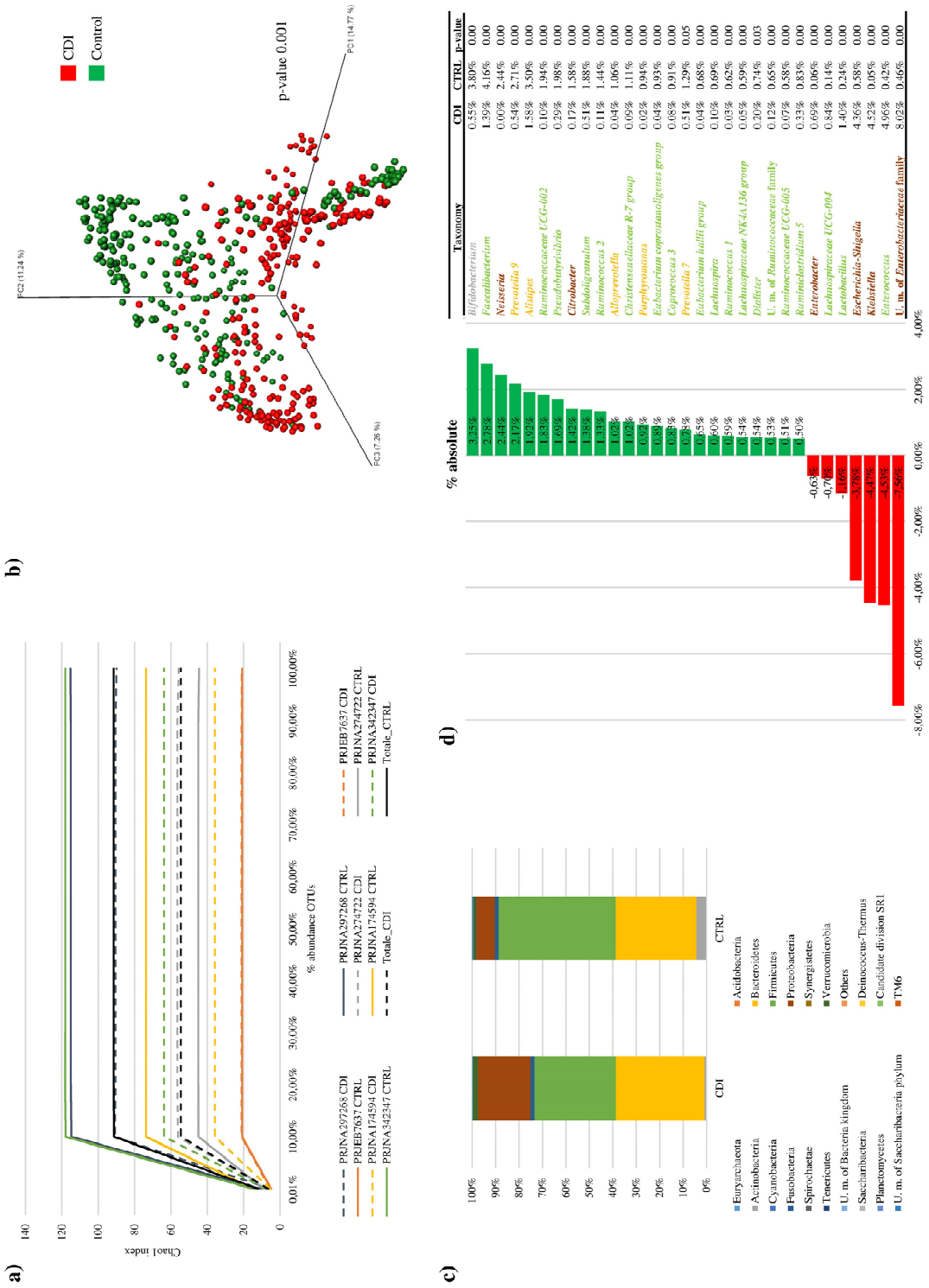


Figure 5. Exploration of the of the microbiota composition of CDI and CTRL samples. Panel a displays the complexity of each studied sample and of CDI and CTRL average calculated through Chao1 and represented with rarefaction curves. Panel b reports the principal coordinate analysis (PCoA) of the collected case-control belonging to CDI studies. Panel c displays the bacterial composition at phylum level based on cross-study of the CDI and CTRL groups. Panel d reports the taxa composition at genus level reporting only bacteria with an average abundance variation > 0.5 %.

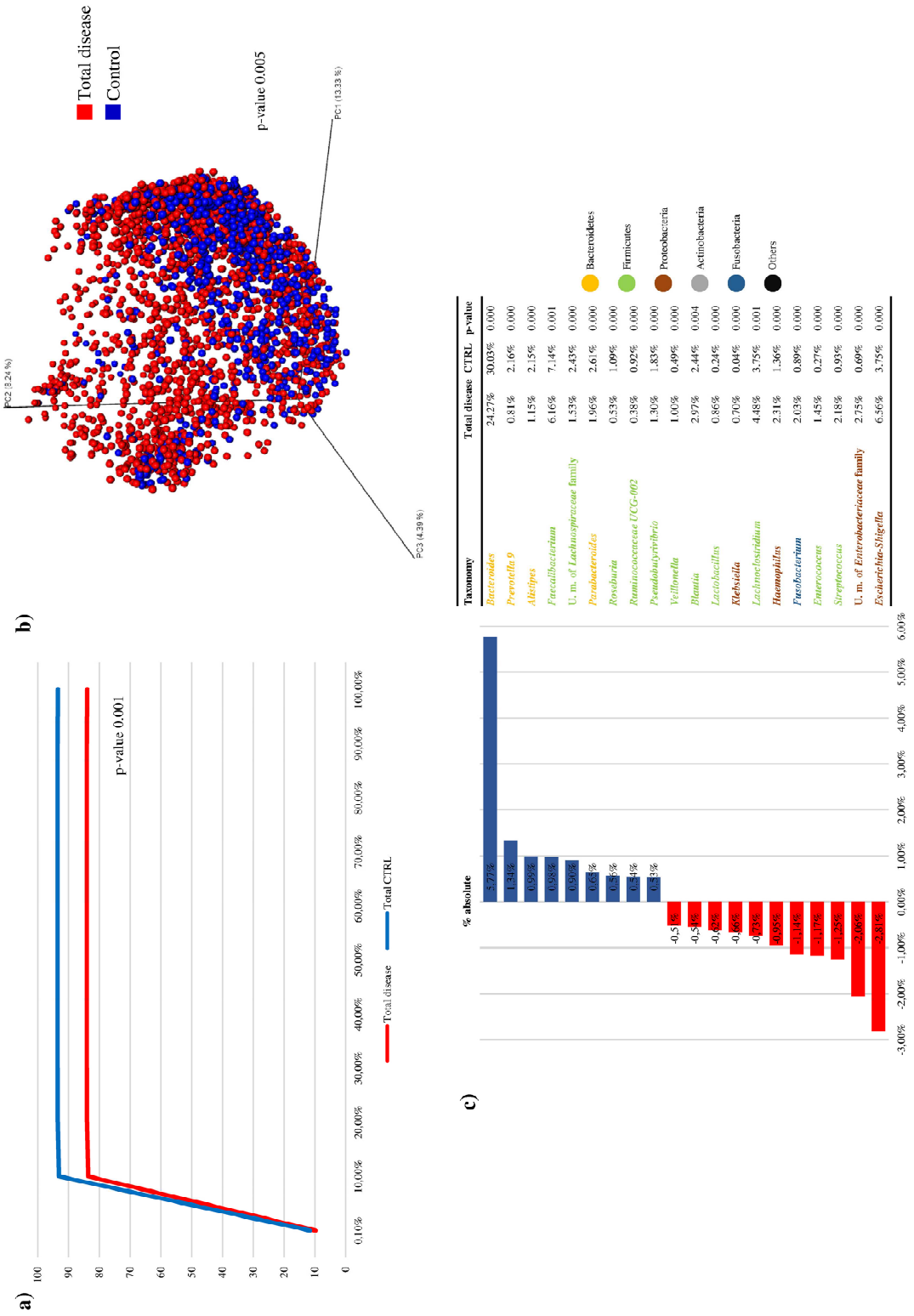


Figure 6. Investigation of the microbiota composition of all subjects affected by an intestinal disease (DS) and control samples. Panel a shows alpha-diversity curves calculated through Chao1 index of each study and of DS and CTRL average. Panel b reports the PCoA of the collected case-control belonging to DS studies. Panel c indicates the bacterial composition at phylum level based on cross-study of the DS and CTRL groups. Panel d shows the bacterial genera that present an average abundance difference > 0.5 %.

Table 1. Summary of the microbial gut biomarkers identified in the study.

Intestinal diseases	Samples	Genera
Crohn's disease	CTRL	*Barnesiella* *Ruminococcus 2*
	CD	*Actinomyces* *Eggerthella* *Blautia* *Peptoclostridium* *Flavonifractor* *Erysipelatoclostridium* *Lactobacillus* *Streptococcus* U. m. of *Proteobacteria* phylum
Ulcerative colitis	CTRL	*Barnesiella* *Odoribacter* *Alistipes*
	UC	*Faecalibacterium* *Streptococcus* *Veillonella* U. m. of *Enterobacteriaceae* family *Haemophilus*
Colorectal cancer	CTRL	U. m. of *Lachnospiraceae* family *Faecalibacterium* *Ruminococcaceae UCG-005* *Subdoligranulum*
	CRC	*Alloprevotella* *Gemella* *Parvimonas* *Streptococcus* *Leptotrichia* *Campylobacter*
Clostridium difficile infection	CTRL	*Christensenellaceae R-7* group U. m. of *Lachnospiraceae* family *Ruminococcaceae UCG-003*
	CDI	*Erysipelatoclostridium* *Enterococcus* *Lactobacillus*
Total diseases	Total CTRL	*Barnesiella* *Ruminococcaceae UCG-005* *Alistipes* *Christensenellaceae R-7* group U. m. of *Lachnospiraceae* family
	Total diseases	*Lactobacillus* *Streptococcus* U. m. of *Enterobacteriaceae* family

Chapter 8

General Conclusion

Reconstruction of chickens' microbiome

The 16S rRNA gene microbial profiling data analysis was performed in order to identify possible differences in cecal samples of chickens from different rearing conditions. In detail, chickens reared under commercial production (Broiler chicken, BC), grown in free roaming conditions (Free-range chicken, FRC) and farm chickens that have undergone a reverse domestication process (Feral chicken, FC) were collected. Taxonomical analysis between the three poultry groups showed that FC microbiota displayed a composition that still resembles that of FRC cecal samples. Interestingly, the reconstructed microbiota of FRC and FC animals were characterized by a higher abundance of strains belonging to the Bacteroidetes phylum, being clearly different from the cecal microbiota of BC domesticated animals. Furthermore, analyzing the cecal microbiota of FC, BC and FRC, 18 taxa belonging to the Firmicutes phylum were present in all samples, thus forming a proposed core microbiota. Analysis of the resistome of cecal microbiomes revealed that BC datasets contain higher levels of predicted antibiotic-resistant genes (ARGs) compared to those of FRC and FC. This observed increase in BC animals indicates that the use of antibiotics may modulates the composition of the cecal microbiota toward antibiotic-resistant bacteria. The prediction of the functional capabilities of the FRC microbiome showed high abundance of genes involved in the pathway for acetate, while BC revealed high presence in genes involved in formate production.

The functional microbiota characterization of FC microbiome revealed a less efficiency in energy recovery compared to BC. Moreover, it has been completed a first attempt to reconstruct the cecal microbiota of FC. Culturomics approaches were applied to *in vitro* reconstruct the microbiota of FC animals, providing comprehensive culture conditions simulating or mimicking the environmental conditions present in the cecum of chickens.

Microbial biomarkers as indicators of human disease

In order to evaluate the correlation between microbiota composition and factors such as diet, lifestyle and environment, it was performed a comparison of publicly available shotgun datasets of human fecal samples collected from different geographical regions (Europe, North America, Asia and Oceania). In detail, it was evaluated the possibility that urbanization/industrialization processes have shaped the gut microbiomes as determined by a meta-analysis of various metagenomic. This analysis allows the prediction of the global human gut core-microbiome, which was shown to encompass 22 genera. The comparison between pre-agricultural vs. urbanized/industrialized gut microbiomes included in the study allowed to identify particular taxa that seem to have been acquired or lost during the urbanization/industrialization process, perhaps as a result of dietary changes. In this context, the increase in dietary diversity and in protein and lipid intake appears to have caused an expansion of the metabolic capabilities of the human gut microbiome towards degradation of a wider range of polysaccharides and utilization of a higher number of available amino acids and lipids. However, the lack of metadata regarding dietary habits and lifestyle of the assessed populations prevents us to make any connection between diet and microbiome composition. Furthermore, antibiotic resistance profiling in the analyzed datasets underlined a progressive increase in AREs proportional to the level of urbanization/industrialization and corresponding intensity of antibiotic treatment.

Subsequently, the correlation between gut microbiota and human intestinal diseases was evaluated. The focus was on functional constipation (FC) that is a widespread gastrointestinal disorder responsible for difficult or infrequent bowel movements. Analysis of gut microbial composition revealed alterations in relative abundance of specific taxa, such as *Bacteroides* and *Feacalibacterium.* It also profiled metabolic pathways of microbiomes corresponding to functionally constipated and healthy individuals. The reconstructed gut microbiome of individuals affected by functional constipation revealed that the FC microbiome is characterized by a high abundance of genes involved in hydrogen production, methanogenesis and glycerol degradation. In contrast, the microbiomes of

HS samples showed an increase of pathways implicated in carbohydrate and fatty acid metabolism, and in methylglyoxal degradation. Alteration of these metabolic pathways appears to impact on functional constipation and related symptoms, thus highlighting the key functional role exerted by the gut microbiome in maintaining the health status of the host.

Moreover, a cross-study meta-analysis of 3048 public metagenomic datasets, encompass control and diseased samples, was performed in order to identify possible bacterial biomarkers for major intestinal diseases such as Crohn's disease (CD), ulcerative colitis (UC), colorectal cancer (CRC) and *Clostridium difficile* infections (CDI). Furthermore, all datasets were analyzed together, in order to identify possible universal gut disease microbial biomarkers. In detail, this cross-study analysis showed that *Barnesiella*, *Ruminococcaceae* UCG-005, *Alistipes*, *Christensenellaceae R-7* group and unclassified member of Lachnospiraceae family genera correlated with a healthy state of subjects. In contrast, subjects that display an intestinal disease displayed higher abundance of genera reported to cause intestinal inflammation, such as unclassified member of Erysipelotrichaceae family and *Streptococcus*. The discovery and the identification of novel universal biomarkers as indicators of health of the human intestinal tract may contribute to quick diagnosis as well as to predict the course and prognosis of the disease and guide therapeutic decisions improving patient care.

References

Allen HK & Stanton TB (2014) Altered egos: antibiotic effects on food animal microbiomes. *Annu Rev Microbiol* 68: 297-315.

Arnal ME, Zhang J, Messori S, Bosi P, Smidt H & Lalles JP (2014) Early changes in microbial colonization selectively modulate intestinal enzymes, but not inducible heat shock proteins in young adult Swine. *PLoS One* 9: e87967.

Baothman OA, Zamzami MA, Taher I, Abubaker J & Abu-Farha M (2016) The role of Gut Microbiota in the development of obesity and Diabetes. *Lipids Health Dis* 15: 108.

Benus RF, van der Werf TS, Welling GW, Judd PA, Taylor MA, Harmsen HJ & Whelan K (2010) Association between Faecalibacterium prausnitzii and dietary fibre in colonic fermentation in healthy human subjects. *Br J Nutr* 104: 693-700.

Berry D & Reinisch W (2013) Intestinal microbiota: a source of novel biomarkers in inflammatory bowel diseases? *Best Pract Res Clin Gastroenterol* 27: 47-58.

Bessegatto JA, Paulino LR, Lisboa JAN, *et al.* (2017) Changes in the fecal microbiota of beef cattle caused by change in management and the use of virginiamycin as a growth promoter. *Res Vet Sci* 114: 355-362.

Biteen JS, Blainey PC, Cardon ZG, *et al.* (2016) Tools for the Microbiome: Nano and Beyond. *Acs Nano* 10: 6-37.

Brauner A, Brandt L, Frisan T, Thelestam M & Ekbom A (2010) Is there a risk of cancer development after Campylobacter infection? *Scand J Gastroenterol* 45: 893-897.

Brinkworth GD, Noakes M, Clifton PM & Bird AR (2009) Comparative effects of very low-carbohydrate, high-fat and high-carbohydrate, low-fat weight-loss diets on bowel habit and faecal short-chain fatty acids and bacterial populations. *British Journal of Nutrition* 101: 1493-1502.

Browne HP, Forster SC, Anonye BO, *et al.* (2016) Culturing of 'unculturable' human microbiota reveals novel taxa and extensive sporulation. *Nature* 533: 543-546.

Cani PD, Neyrinck AM, Tuohy KM, *et al.* (2007) Changes in gut microflora are responsible for high-fat diet-induced diabetes through a mechanism associated with endotoxaemia. *Diabetologia* 50: S68-S68.

Carroll IM, Ringel-Kulka T, Keku TO, Chang YH, Packey CD, Sartor RB & Ringel Y (2011) Molecular analysis of the luminal- and mucosal-associated intestinal microbiota in diarrhea-predominant irritable bowel syndrome. *Am J Physiol Gastrointest Liver Physiol* 301: G799-807.

Carvalho-Wells AL, Helmolz K, Nodet C, *et al.* (2010) Determination of the in vivo prebiotic potential of a maize-based whole grain breakfast cereal: a human feeding study. *Br J Nutr* 104: 1353-1356.

Castellarin M, Warren RL, Freeman JD, *et al.* (2012) Fusobacterium nucleatum infection is prevalent in human colorectal carcinoma. *Genome Res* 22: 299-306.

Chakravorty S, Helb D, Burday M, Connell N & Alland D (2007) A detailed analysis of 16S ribosomal RNA gene segments for the diagnosis of pathogenic bacteria. *Journal of Microbiological Methods* 69: 330-339.

Choi KY, Lee TK & Sul WJ (2015) Metagenomic Analysis of Chicken Gut Microbiota for Improving Metabolism and Health of Chickens - A Review. *Asian-Australas J Anim Sci* 28: 1217-1225.

Cisek AA & Binek M (2014) Chicken intestinal microbiota function with a special emphasis on the role of probiotic bacteria. *Pol J Vet Sci* 17: 385-394.

Claesson MJ, Wang QO, O'Sullivan O, Greene-Diniz R, Cole JR, Ross RP & O'Toole PW (2010) Comparison of two next-generation sequencing technologies for resolving highly

complex microbiota composition using tandem variable 16S rRNA gene regions. *Nucleic Acids Research* **38.**
Clarridge JE (2004) Impact of 16S rRNA gene sequence analysis for identification of bacteria on clinical microbiology and infectious diseases. *Clinical Microbiology Reviews* **17: 840-+.**
Conde L, Halperin E, Akers NK, *et al.* (2010) Genome-wide association study of follicular lymphoma identifies a risk locus at 6p21.32. *Nat Genet* **42: 661-664.**
Costabile A, Kolida S, Klinder A, *et al.* (2010) A double-blind, placebo-controlled, cross-over study to establish the bifidogenic effect of a very-long-chain inulin extracted from globe artichoke (Cynara scolymus) in healthy human subjects. *Br J Nutr* **104: 1007-1017.**
Costello EK, Stagaman K, Dethlefsen L, Bohannan BJ & Relman DA (2012) The application of ecological theory toward an understanding of the human microbiome. *Science* **336: 1255-1262.**
David LA, Maurice CF, Carmody RN, *et al.* (2014) Diet rapidly and reproducibly alters the human gut microbiome. *Nature* **505: 559-563.**
Davis CD & Milner JA (2009) Gastrointestinal microflora, food components and colon cancer prevention. *J Nutr Biochem* **20: 743-752.**
De Filippo C, Cavalieri D, Di Paola M, *et al.* (2010) Impact of diet in shaping gut microbiota revealed by a comparative study in children from Europe and rural Africa. *Proc Natl Acad Sci U S A* **107: 14691-14696.**
Deusch S, Tilocca B, Camarinha-Silva A & Seifert J (2015) News in livestock research - use of Omics-technologies to study the microbiota in the gastrointestinal tract of farm animals. *Comput Struct Biotechnol J* **13: 55-63.**
Dicksved J, Halfvarson J, Rosenquist M, *et al.* (2008) Molecular analysis of the gut microbiota of identical twins with Crohn's disease. *ISME J* **2: 716-727.**
Dominguez-Bello MG, Costello EK, Contreras M, Magris M, Hidalgo G, Fierer N & Knight R (2010) Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns. *Proc Natl Acad Sci U S A* **107: 11971-11975.**
Duncan SH, Belenguer A, Holtrop G, Johnstone AM, Flint HJ & Lobley GE (2007) Reduced dietary intake of carbohydrates by obese subjects results in decreased concentrations of butyrate and butyrate-producing bacteria in feces. *Applied and Environmental Microbiology* **73: 1073-1078.**
Eggesbo M, Botten G, Stigum H, Nafstad P & Magnus P (2003) Is delivery by cesarean section a risk factor for food allergy? *Journal of Allergy and Clinical Immunology* **112: 420-426.**
Favier C, Neut C, Mizon C, Cortot A, Colombel JF & Mizon J (1997) Fecal beta-D-galactosidase production and Bifidobacteria are decreased in Crohn's disease. *Dig Dis Sci* **42: 817-822.**
Fecteau ME, Pitta DW, Vecchiarelli B, *et al.* (2016) Dysbiosis of the Fecal Microbiota in Cattle Infected with Mycobacterium avium subsp. paratuberculosis. *PLoS One* **11: e0160353.**
Gagniere J, Raisch J, Veziant J, *et al.* (2016) Gut microbiota imbalance and colorectal cancer. *World J Gastroenterol* **22: 501-518.**
Gao R, Gao Z, Huang L & Qin H (2017) Gut microbiota and colorectal cancer. *Eur J Clin Microbiol Infect Dis* **36: 757-769.**
Garrett WS, Gallini CA, Yatsunenko T, *et al.* (2010) Enterobacteriaceae act in concert with the gut microbiota to induce spontaneous and maternally transmitted colitis. *Cell Host Microbe* **8: 292-300.**
Gerritsen J, Smidt H, Rijkers GT & de Vos WM (2011) Intestinal microbiota in human health and disease: the impact of probiotics. *Genes Nutr* **6: 209-240.**
Goldin BR & Gorbach SL (1980) Effect of Lactobacillus acidophilus dietary supplements on 1,2-dimethylhydrazine dihydrochloride-induced intestinal cancer in rats. *J Natl Cancer Inst* **64: 263-265.**
Goldin BR, Gualtieri LJ & Moore RP (1996) The effect of Lactobacillus GG on the initiation and promotion of DMH-induced intestinal tumors in the rat. *Nutr Cancer* **25: 197-204.**

Gomez DE, Arroyo LG, Costa MC, Viel L & Weese JS (2017) Characterization of the Fecal Bacterial Microbiota of Healthy and Diarrheic Dairy Calves. *J Vet Intern Med* 31: 928-939.
Gong J, Si W, Forster RJ*, et al.* (2007) 16S rRNA gene-based analysis of mucosa-associated bacterial community and phylogeny in the chicken gastrointestinal tracts: from crops to ceca. *FEMS Microbiol Ecol* 59: 147-157.
Goodman AL, Kallstrom G, Faith JJ, Reyes A, Moore A, Dantas G & Gordon JI (2011) Extensive personal human gut microbiota culture collections characterized and manipulated in gnotobiotic mice. *Proc Natl Acad Sci U S A* 108: 6252-6257.
Grivennikov SI, Wang K, Mucida D*, et al.* (2012) Adenoma-linked barrier defects and microbial products drive IL-23/IL-17-mediated tumour growth. *Nature* 491: 254-258.
Guinane CM & Cotter PD (2013) Role of the gut microbiota in health and chronic gastrointestinal disease: understanding a hidden metabolic organ. *Therap Adv Gastroenterol* 6: 295-308.
Hallert C, Bjorck I, Nyman M, Pousette A, Granno C & Svensson H (2003) Increasing fecal butyrate in ulcerative colitis patients by diet: controlled pilot study. *Inflamm Bowel Dis* 9: 116-121.
Hamady M & Knight R (2009) Microbial community profiling for human microbiome projects: Tools, techniques, and challenges. *Genome Res* 19: 1141-1152.
Hawrelak JA & Myers SP (2004) The causes of intestinal dysbiosis: a review. *Altern Med Rev* 9: 180-197.
Heinritz SN, Weiss E, Eklund M*, et al.* (2016) Intestinal Microbiota and Microbial Metabolites Are Changed in a Pig Model Fed a High-Fat/Low-Fiber or a Low-Fat/High-Fiber Diet. *PLoS One* 11: e0154329.
Hugon P, Dufour JC, Colson P, Fournier PE, Sallah K & Raoult D (2015) A comprehensive repertoire of prokaryotic species identified in human beings. *Lancet Infect Dis* 15: 1211-1219.
Huh SY, Rifas-Shiman SL, Zera CA, Edwards JW, Oken E, Weiss ST & Gillman MW (2012) Delivery by caesarean section and risk of obesity in preschool age children: a prospective cohort study. *Arch Dis Child* 97: 610-616.
Kabeerdoss J, Devi RS, Mary RR & Ramakrishna BS (2012) Faecal microbiota composition in vegetarians: comparison with omnivores in a cohort of young women in southern India. *Br J Nutr* 108: 953-957.
Kankaanpaa PE, Salminen SJ, Isolauri E & Lee YK (2001) The influence of polyunsaturated fatty acids on probiotic growth and adhesion. *Fems Microbiology Letters* 194: 149-153.
Kerr B, Riley MA, Feldman MW & Bohannan BJ (2002) Local dispersal promotes biodiversity in a real-life game of rock-paper-scissors. *Nature* 418: 171-174.
Kim A (2015) Dysbiosis: A Review Highlighting Obesity and Inflammatory Bowel Disease. *J Clin Gastroenterol* 49 Suppl 1: S20-24.
Kim HB, Borewicz K, White BA, Singer RS, Sreevatsan S, Tu ZJ & Isaacson RE (2012) Microbial shifts in the swine distal gut in response to the treatment with antimicrobial growth promoter, tylosin. *Proc Natl Acad Sci U S A* 109: 15485-15490.
Koren O, Goodrich JK, Cullender TC*, et al.* (2012) Host remodeling of the gut microbiome and metabolic changes during pregnancy. *Cell* 150: 470-480.
Lagier JC, Armougom F, Million M*, et al.* (2012) Microbial culturomics: paradigm shift in the human gut microbiome study. *Clin Microbiol Infect* 18: 1185-1193.
Lagier JC, Khelaifia S, Alou MT*, et al.* (2016) Culture of previously uncultured members of the human gut microbiota by culturomics. *Nat Microbiol* 1: 16203.
Ley RE, Turnbaugh PJ, Klein S & Gordon JI (2006) Microbial ecology: human gut microbes associated with obesity. *Nature* 444: 1022-1023.
Ley RE, Lozupone CA, Hamady M, Knight R & Gordon JI (2008) Worlds within worlds: evolution of the vertebrate gut microbiota. *Nat Rev Microbiol* 6: 776-788.

Li Q, Wang C, Tang C, Li N & Li J (2012) Molecular-phylogenetic characterization of the microbiota in ulcerated and non-ulcerated regions in the patients with Crohn's disease. *PLoS One* 7: e34939.
Liang Q, Chiu J, Chen Y, *et al.* (2017) Fecal Bacteria Act as Novel Biomarkers for Noninvasive Diagnosis of Colorectal Cancer. *Clin Cancer Res* 23: 2061-2070.
Looft T, Allen HK, Cantarel BL, *et al.* (2014) Bacteria, phages and pigs: the effects of in-feed antibiotics on the microbiome at different gut locations. *ISME J* 8: 1566-1576.
Lozupone CA, Stombaugh JI, Gordon JI, Jansson JK & Knight R (2012) Diversity, stability and resilience of the human gut microbiota. *Nature* 489: 220-230.
Lugli GA, Milani C, Mancabelli L, *et al.* (2017) Erratum to: Ancient bacteria of the Otzi's microbiome: a genomic tale from the Copper Age. *Microbiome* 5: 23.
Mai V, Young CM, Ukhanova M, *et al.* (2011) Fecal microbiota in premature infants prior to necrotizing enterocolitis. *PLoS One* 6: e20647.
Mangin I, Bonnet R, Seksik P, *et al.* (2004) Molecular inventory of faecal microflora in patients with Crohn's disease. *FEMS Microbiol Ecol* 50: 25-36.
Manichanh C, Rigottier-Gois L, Bonnaud E, *et al.* (2006) Reduced diversity of faecal microbiota in Crohn's disease revealed by a metagenomic approach. *Gut* 55: 205-211.
Mao S, Zhang M, Liu J & Zhu W (2015) Characterising the bacterial microbiota across the gastrointestinal tracts of dairy cattle: membership and potential function. *Sci Rep* 5: 16116.
Marchesi JR, Holmes E, Khan F, *et al.* (2007) Rapid and noninvasive metabonomic characterization of inflammatory bowel disease. *J Proteome Res* 6: 546-551.
Maslowski KM, Vieira AT, Ng A, *et al.* (2009) Regulation of inflammatory responses by gut microbiota and chemoattractant receptor GPR43. *Nature* 461: 1282-1286.
Matijasic BB, Obermajer T, Lipoglavsek L, Grabnar I, Avgustin G & Rogelj I (2014) Association of dietary type with fecal microbiota in vegetarians and omnivores in Slovenia. *Eur J Nutr* 53: 1051-1064.
Matsuoka K & Kanai T (2015) The gut microbiota and inflammatory bowel disease. *Semin Immunopathol* 37: 47-55.
Milani C, Turroni F, Duranti S, *et al.* (2015) Genomics of the Genus Bifidobacterium Reveals Species-Specific Adaptation to the Glycan-Rich Gut Environment. *Appl Environ Microbiol* 82: 980-991.
Milani C, Lugli GA, Turroni F, *et al.* (2014) Evaluation of bifidobacterial community composition in the human gut by means of a targeted amplicon sequencing (ITS) protocol. *FEMS Microbiol Ecol* 90: 493-503.
Milani C, Hevia A, Foroni E, *et al.* (2013) Assessing the fecal microbiota: an optimized ion torrent 16S rRNA gene-based analysis protocol. *PLoS One* 8: e68739.
Milani C, Ticinesi A, Gerritsen J, *et al.* (2016) Gut microbiota composition and Clostridium difficile infection in hospitalized elderly individuals: a metagenomic study. *Sci Rep* 6: 25945.
Mima K, Sukawa Y, Nishihara R, *et al.* (2015) Fusobacterium nucleatum and T Cells in Colorectal Carcinoma. *JAMA Oncol* 1: 653-661.
Neefs JM, Vandepeer Y, Derijk P, Chapelle S & Dewachter R (1993) Compilation of Small Ribosomal-Subunit Rna Structures. *Nucleic Acids Research* 21: 3025-3049.
Nemoto H, Kataoka K, Ishikawa H, *et al.* (2012) Reduced diversity and imbalance of fecal microbiota in patients with ulcerative colitis. *Dig Dis Sci* 57: 2955-2964.
Nicholson JK, Holmes E, Kinross J, Burcelin R, Gibson G, Jia W & Pettersson S (2012) Host-gut microbiota metabolic interactions. *Science* 336: 1262-1267.
Niu Q, Li P, Hao S, *et al.* (2015) Dynamic distribution of the gut microbiota and the relationship with apparent crude fiber digestibility and growth stages in pigs. *Sci Rep* 5: 9938.
O'Keefe SJ, Ou J, Aufreiter S, *et al.* (2009) Products of the colonic microbiota mediate the effects of diet on colon cancer risk. *J Nutr* 139: 2044-2048.

Ou J, Carbonero F, Zoetendal EG, *et al.* (2013) Diet, microbiota, and microbial metabolites in colon cancer risk in rural Africans and African Americans. *Am J Clin Nutr* 98: 111-120.
Pokusaeva K, Fitzgerald GF & van Sinderen D (2011) Carbohydrate metabolism in Bifidobacteria. *Genes Nutr* 6: 285-306.
Pourabedin M & Zhao X (2015) Prebiotics and gut microbiota in chickens. *FEMS Microbiol Lett* 362: fnv122.
Rajilic-Stojanovic M & de Vos WM (2014) The first 1000 cultured species of the human gastrointestinal microbiota. *Fems Microbiology Reviews* 38: 996-1047.
Rinke C, Schwientek P, Sczyrba A, *et al.* (2013) Insights into the phylogeny and coding potential of microbial dark matter. *Nature* 499: 431-437.
Rowland IR, Rumney CJ, Coutts JT & Lievense LC (1998) Effect of Bifidobacterium longum and inulin on gut bacterial metabolism and carcinogen-induced aberrant crypt foci in rats. *Carcinogenesis* 19: 281-285.
Rubinstein MR, Wang X, Liu W, Hao Y, Cai G & Han YW (2013) Fusobacterium nucleatum promotes colorectal carcinogenesis by modulating E-cadherin/beta-catenin signaling via its FadA adhesin. *Cell Host Microbe* 14: 195-206.
Sabbioni A, Ferrario C, Milani C, *et al.* (2016) Modulation of the Bifidobacterial Communities of the Dog Microbiota by Zeolite. *Front Microbiol* 7: 1491.
Salminen S, Gibson GR, McCartney AL & Isolauri E (2004) Influence of mode of delivery on gut microbiota composition in seven year old children. *Gut* 53: 1388-1389.
Sanderson IR (2004) Short chain fatty acid regulation of signaling genes expressed by the intestinal epithelium. *J Nutr* 134: 2450S-2454S.
Saulnier DM, Riehle K, Mistretta TA, *et al.* (2011) Gastrointestinal microbiome signatures of pediatric patients with irritable bowel syndrome. *Gastroenterology* 141: 1782-1791.
Scheppach W, Sommer H, Kirchner T, *et al.* (1992) Effect of butyrate enemas on the colonic mucosa in distal ulcerative colitis. *Gastroenterology* 103: 51-56.
Schloss PD, Westcott SL, Ryabin T, *et al.* (2009) Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl Environ Microbiol* 75: 7537-7541.
Schnorr SL, Candela M, Rampelli S, *et al.* (2014) Gut microbiome of the Hadza hunter-gatherers. *Nat Commun* 5: 3654.
Segain JP, Raingeard de la Bletiere D, Bourreille A, *et al.* (2000) Butyrate inhibits inflammatory responses through NFkappaB inhibition: implications for Crohn's disease. *Gut* 47: 397-403.
Segata N, Waldron L, Ballarini A, Narasimhan V, Jousson O & Huttenhower C (2012) Metagenomic microbial community profiling using unique clade-specific marker genes. *Nat Methods* 9: 811-814.
Sergeant MJ, Constantinidou C, Cogan TA, Bedford MR, Penn CW & Pallen MJ (2014) Extensive microbial and functional diversity within the chicken cecal microbiome. *PLoS One* 9: e91941.
Sevelsted A, Stokholm J, Bonnelykke K & Bisgaard H (2015) Cesarean section and chronic immune disorders. *Pediatrics* 135: e92-98.
Smillie CS, Smith MB, Friedman J, Cordero OX, David LA & Alm EJ (2011) Ecology drives a global network of gene exchange connecting the human microbiome. *Nature* 480: 241-244.
Sobhani I, Tap J, Roudot-Thoraval F, *et al.* (2011) Microbial Dysbiosis in Colorectal Cancer (CRC) Patients. *Plos One* 6.
Sokol H, Pigneur B, Watterlot L, *et al.* (2008) Faecalibacterium prausnitzii is an anti-inflammatory commensal bacterium identified by gut microbiota analysis of Crohn disease patients. *Proc Natl Acad Sci U S A* 105: 16731-16736.
Song M, Xia B & Li J (2006) Effects of topical treatment of sodium butyrate and 5-aminosalicylic acid on expression of trefoil factor 3, interleukin 1beta, and nuclear factor kappaB in trinitrobenzene sulphonic acid induced colitis in rats. *Postgrad Med J* 82: 130-135.

Spor A, Koren O & Ley R (2011) Unravelling the effects of the environment and host genotype on the gut microbiome. ***Nat Rev Microbiol*** **9: 279-290.**
Stanley D, Geier MS, Denman SE, Haring VR, Crowley TM, Hughes RJ & Moore RJ (2013) Identification of chicken intestinal microbiota correlated with the efficiency of energy extraction from feed. ***Vet Microbiol*** **164: 85-92.**
Sundquist A, Bigdeli S, Jalili R, ***et al.*** **(2007) Bacterial flora-typing with targeted, chip-based Pyrosequencing.** ***Bmc Microbiology*** **7.**
Tamboli CP, Neut C, Desreumaux P & Colombel JF (2004) Dysbiosis in inflammatory bowel disease. ***Gut*** **53: 1-4.**
Turnbaugh PJ, Ley RE, Mahowald MA, Magrini V, Mardis ER & Gordon JI (2006) An obesity-associated gut microbiome with increased capacity for energy harvest. ***Nature*** **444: 1027-1031.**
Turnbaugh PJ, Ridaura VK, Faith JJ, Rey FE, Knight R & Gordon JI (2009) The effect of diet on the human gut microbiome: a metagenomic analysis in humanized gnotobiotic mice. ***Sci Transl Med*** **1: 6ra14.**
Turroni F, Peano C, Pass DA, ***et al.*** **(2012) Diversity of bifidobacteria within the infant gut microbiota.** ***PLoS One*** **7: e36957.**
Uronis JM & Threadgill DW (2009) Murine models of colorectal cancer. ***Mamm Genome*** **20: 261-268.**
Uronis JM, Muhlbauer M, Herfarth HH, Rubinas TC, Jones GS & Jobin C (2009) Modulation of the intestinal microbiota alters colitis-associated colorectal cancer susceptibility. ***PLoS One*** **4: e6026.**
Ursell LK, Clemente JC, Rideout JR, Gevers D, Caporaso JG & Knight R (2012) The interpersonal and intrapersonal diversity of human-associated microbiota in key body sites. ***J Allergy Clin Immunol*** **129: 1204-1208.**
van der Wielen PW, Keuzenkamp DA, Lipman LJ, van Knapen F & Biesterveld S (2002) Spatial and temporal variation of the intestinal bacterial community in commercially raised broiler chickens during growth. ***Microb Ecol*** **44: 286-293.**
van Nuenen MH, Venema K, van der Woude JC & Kuipers EJ (2004) The metabolic activity of fecal microbiota from healthy individuals and patients with inflammatory bowel disease. ***Dig Dis Sci*** **49: 485-491.**
Ventura M, Turroni F, Canchaya C, Vaughan EE, O'Toole PW & van Sinderen D (2009) Microbial diversity in the human intestine and novel insights from metagenomics. ***Front Biosci (Landmark Ed)*** **14: 3214-3221.**
Ventura M, O'Flaherty S, Claesson MJ, Turroni F, Klaenhammer TR, van Sinderen D & O'Toole PW (2009) Genome-scale analyses of health-promoting bacteria: probiogenomics. ***Nat Rev Microbiol*** **7: 61-71.**
Walker AW, Duncan SH, Louis P & Flint HJ (2014) Phylogeny, culturing, and metagenomics of the human gut microbiota. ***Trends Microbiol*** **22: 267-274.**
Walker AW, Duncan SH, McWilliam Leitch EC, Child MW & Flint HJ (2005) pH and peptide supply can radically alter bacterial populations and short-chain fatty acid ratios within microbial communities from the human colon. ***Appl Environ Microbiol*** **71: 3692-3700.**
Wang L, Lilburn M & Yu Z (2016) Intestinal Microbiota of Broiler Chickens As Affected by Litter Management Regimens. ***Front Microbiol*** **7: 593.**
Wei Z, Cao S, Liu S, ***et al.*** **(2016) Could gut microbiota serve as prognostic biomarker associated with colorectal cancer patients' survival? A pilot study on relevant mechanism.** ***Oncotarget*** **7: 46158-46172.**
Wong JM, de Souza R, Kendall CW, Emam A & Jenkins DJ (2006) Colonic health: fermentation and short chain fatty acids. ***J Clin Gastroenterol*** **40: 235-243.**
Wu M, McNulty NP, Rodionov DA, ***et al.*** **(2015) Genetic determinants of in vivo fitness and diet responsiveness in multiple human gut Bacteroides.** ***Science*** **350: aac5992.**

Yatsunenko T, Rey FE, Manary MJ, *et al.* (2012) Human gut microbiome viewed across age and geography. *Nature* 486: 222-227.
Yu T, Guo F, Yu Y, *et al.* (2017) Fusobacterium nucleatum Promotes Chemoresistance to Colorectal Cancer by Modulating Autophagy. *Cell* 170: 548-563 e516.
Zhou W, Wang Y & Lin J (2012) Functional cloning and characterization of antibiotic resistance genes from the chicken gut microbiome. *Appl Environ Microbiol* 78: 3028-3032.

Publications

1. Turroni F, Milani C, Duranti S, Ferrario C, Lugli GA, Mancabelli L, van Sinderen D, Ventura M. **"Bifidobacteria and the infant gut: an example of co-evolution and natural selection"** Cell Mol Life Sci. 2017 Oct 5. doi: 10.1007/s00018-017-2672-0.

2. Ferrario C, Alessandri G, **Mancabelli L**, Gering E, Mangifesta M, Milani C, Lugli GA, Viappiani A, Duranti S, Turroni F, Ossiprandi MC, Hiyashi R, Mackie R, van Sinderen D, Ventura M. **"Untangling the cecal microbiota of feral chickens by culturomic and metagenomic analyses"** Environ Microbiol. 2017 Oct 2. doi: 10.1111/1462-2920.13943.

3. Ferrario C, Statello R, Carnevali L, **Mancabelli L**, Milani C, Mangifesta M, Duranti S, Lugli GA, Jimenez B, Lodge S, Viappiani A, Alessandri G, Dall'Asta M, Del Rio D, Sgoifo A, van Sinderen D, Ventura M, Turroni F. **"How to Feed the Mammalian Gut Microbiota: Bacterial and Metabolic Modulation by Dietary Fibers"** Front. Microbiol., 12 September 2017 | https://doi.org/10.3389/fmicb.2017.01749.

4. Ticinesi A, Milani C, Lauretani F, Nouvenne A, **Mancabelli L**, Lugli GA, Turroni F, Duranti S, Mangifesta M, Viappiani A, Ferrario C, Maggio M, Ventura M, Meschi T. **"Gut microbiota composition is associated with polypharmacy in elderly hospitalized patients."** Sci Rep. 2017 Sep 11;7(1):11102. doi: 10.1038/s41598-017-10734-y.

5. Duranti S, Mangifesta M, Lugli GA, Turroni F, Anzalone R, Milani C, **Mancabelli L**, Ossiprandi MC, Ventura M. **"*Bifidobacterium vansinderenii* sp. nov., isolated from faeces of emperor tamarin (*Saguinus imperator*)."** Int J Syst Evol Microbiol. 2017 Sep 12. doi: 10.1099/ijsem.0.002243.

6. **Mancabelli L**, Milani C, Lugli GA, Turroni F, Mangifesta M, Viappiani A, Ticinesi A, Nouvenne A, Meschi T, van Sinderen D, Ventura M. **"Unveiling the gut microbiota composition and functionality associated with constipation through metagenomic analyses."** Sci Rep. 2017 Aug 29;7(1):9879. doi: 10.1038/s41598-017-10663-w.

7. Milani C, Mangifesta M, **Mancabelli L**, Lugli GA, James K, Duranti S, Turroni F, Ferrario C, Ossiprandi MC, van Sinderen D, Ventura M. **"Unveiling bifidobacterial biogeography across the mammalian branch of the tree of life."** ISME J. 2017 Aug 22. doi: 10.1038/ismej.2017.138.

8. Lugli GA, Milani C, Turroni F, Duranti S, **Mancabelli L**, Mangifesta M, Ferrario C, Modesto M, Mattarelli P, Jiří K, van Sinderen D, Ventura M. **"Comparative genomic and phylogenomic analyses of the *Bifidobacteriaceae* family."** BMC Genomics. 2017 Aug 1;18(1):568. doi: 10.1186/s12864-017-3955-4.

9. Milani C, Mangifesta M, **Mancabelli L**, Lugli GA, Mancino W, Viappiani A, Faccini A, van Sinderen D, Ventura M, Turroni F. **"The Sortase-Dependent Fimbriome of the Genus Bifidobacterium: Extracellular Structures with Potential To Modulate Microbe-Host Dialogue."** Appl Environ Microbiol. 2017 Sep 15;83(19). pii: e01295-17. doi: 10.1128/AEM.01295-17. Print 2017 Oct 1.

10. Duranti S, Lugli GA, **Mancabelli L**, Armanini F, Turroni F, James K, Ferretti P, Gorfer V, Ferrario C, Milani C, Mangifesta M, Anzalone R, Zolfo M, Viappiani A, Pasolli E, Bariletti

I, Canto R, Clementi R, Cologna M, Crifò T, Cusumano G, Fedi S, Gottardi S, Innamorati C, Masè C, Postai D, Savoi D, Soffiati M, Tateo S, Pedrotti A, Segata N, van Sinderen D, Ventura M. **"Maternal inheritance of bifidobacterial communities and bifidophages in infants through vertical transmission."** Microbiome. 2017 Jun 26;5(1):66. doi: 10.1186/s40168-017-0282-6.

11. Ferrario C, Lugli GA, Ossiprandi MC, Turroni F, Milani C, Duranti S, **Mancabelli L**, Mangifesta M, Alessandri G, van Sinderen D, Ventura M. **"Next generation sequencing-based multigene panel for high throughput detection of food-borne pathogens."** Int J Food Microbiol. 2017 Sep 1;256:20-29. doi: 10.1016/j.ijfoodmicro.2017.05.001. Epub 2017 May 3.

12. **Mancabelli L**, Milani C, Lugli GA, Turroni F, Ferrario C, van Sinderen D, Ventura M. **"Meta-analysis of the human gut microbiome from urbanized and pre-agricultural populations."** Environ Microbiol. 2017 Apr;19(4):1379-1390. doi: 10.1111/1462-2920.13692. Epub 2017 Feb 28. Review.

13. Lugli GA, Milani C, **Mancabelli L**, Turroni F, Ferrario C, Duranti S, van Sinderen D, Ventura M. **"Ancient bacteria of the Ötzi's microbiome: a genomic tale from the Copper Age."** Microbiome. 2017 Jan 17;5(1):5. doi: 10.1186/s40168-016-0221-y. Erratum in: Microbiome. 2017 Feb 17;5(1):23.

14. Duranti S, Lugli GA, **Mancabelli L**, Turroni F, Milani C, Mangifesta M, Ferrario C, Anzalone R, Viappiani A, van Sinderen D, Ventura M. **"Prevalence of Antibiotic Resistance Genes among Human Gut-Derived Bifidobacteria."** Appl Environ Microbiol. 2017 Jan 17;83(3). pii: e02894-16. doi: 10.1128/AEM.02894-16. Print 2017 Feb 1.

15. Sabbioni A, Ferrario C, Milani C, **Mancabelli L**, Riccardi E, Di Ianni F, Beretti V, Superchi P, Ossiprandi MC. **"Modulation of the Bifidobacterial Communities of the Dog Microbiota by Zeolite."** Front Microbiol. 2016 Sep 22;7:1491.

16. Duranti S, Gaiani F, **Mancabelli L**, Milani C, Grandi A, Bolchi A, Santoni A, Lugli GA, Ferrario C, Mangifesta M, Viappiani A, Bertoni S, Vivo V, Serafini F, Barbaro MR, Fugazza A, Barbara G, Gioiosa L, Palanza P, Cantoni AM, de'Angelis GL, Barocelli E, de'Angelis N, van Sinderen D, Ventura M, Turroni F. **"Elucidating the gut microbiome of ulcerative colitis: bifidobacteria as novel microbial biomarkers."** FEMS Microbiol Ecol. 2016 Dec;92(12). pii: fiw191. doi: 10.1093/femsec/fiw191. Epub 2016 Sep 7.

17. Milani C, Ticinesi A, Gerritsen J, Nouvenne A, Lugli GA, **Mancabelli L**, Turroni F, Duranti S, Mangifesta M, Viappiani A, Ferrario C, Maggio M, Lauretani F, De Vos W, van Sinderen D, Meschi T, Ventura M. **"Gut microbiota composition and *Clostridium difficile* infection in hospitalized elderly individuals: a metagenomic study."** Sci Rep. 2016 May 11;6:25945. doi: 10.1038/srep25945.

18. **Mancabelli L**, Ferrario C, Milani C, Mangifesta M, Turroni F, Duranti S, Lugli GA, Viappiani A, Ossiprandi MC, van Sinderen D, Ventura M. **"Insights into the biodiversity of the gut microbiota of broiler chickens."** Environ Microbiol. 2016 Apr 30. doi: 10.1111/1462-2920.13363. [Epub ahead of print].

19. Duranti S, Milani C, Lugli GA, **Mancabelli L**, Turroni F, Ferrario C, Mangifesta M, Viappiani A, Sánchez B, Margolles A, van Sinderen D, Ventura M. **"Evaluation of genetic**

diversity among strains of the human gut commensal *Bifidobacterium adolescentis*." Sci Rep. 2016 Apr 1;6:23971. doi: 10.1038/srep23971.

20. Ferrario C, Milani C, **Mancabelli L**, Lugli GA, Duranti S, Mangifesta M, Viappiani A, Turroni F, Margolles A, Ruas-Madiedo P, van Sinderen D, Ventura M. **"Modulation of the eps-ome transcription of bifidobacteria through simulation of human intestinal environment."** FEMS Microbiol Ecol. 2016 Apr;92(4):fiw056. doi: 10.1093/femsec/fiw056. Epub 2016 Mar 8.

21. Lugli GA, Milani C, **Mancabelli L**, van Sinderen D, Ventura M. **"MEGAnnotator: a user-friendly pipeline for microbial genomes assembly and annotation."** FEMS Microbiol Lett. 2016 Apr;363(7). pii: fnw049. doi: 10.1093/femsle/fnw049. Epub 2016 Mar 1.

22. Turroni F, Milani C, Duranti S, **Mancabelli L**, Mangifesta M, Viappiani A, Lugli GA, Ferrario C, Gioiosa L, Ferrarini A, Li J, Palanza P, Delledonne M, van Sinderen D, Ventura M. **"Deciphering bifidobacterial-mediated metabolic interactions and their impact on gut microbiota by a multi-omics approach."** ISME J. 2016 Jul;10(7):1656-68. doi: 10.1038/ismej.2015.236. Epub 2016 Feb 9.

23. Ferrario C, Duranti S, Milani C, **Mancabelli L**, Lugli GA, Turroni F, Mangifesta M, Viappiani A, Ossiprandi MC, van Sinderen D, Ventura M. **"Exploring Amino Acid Auxotrophy in *Bifidobacterium bifidum* PRL2010."** Front Microbiol. 2015 Nov 24;6:1331. doi: 10.3389/fmicb.2015.01331. eCollection 2015.

24. Lugli GA, Milani C, Turroni F, Tremblay D, Ferrario C, **Mancabelli L**, Duranti S, Ward DV, Ossiprandi MC, Moineau S, van Sinderen D, Ventura M. **"Prophages of the genus *Bifidobacterium* as modulating agents of the infant gut microbiota."** Environ Microbiol. 2016 Jul;18(7):2196-213. doi: 10.1111/1462-2920.13154. Epub 2016 Jan 18.

25. Milani C, Turroni F, Duranti S, Lugli GA, **Mancabelli L**, Ferrario C, van Sinderen D, Ventura M. **"Genomics of the Genus *Bifidobacterium* Reveals Species-Specific Adaptation to the Glycan-Rich Gut Environment."** Appl Environ Microbiol. 2015 Nov 20;82(4):980-91. doi: 10.1128/AEM.03500-15. Review.

26. Milani C, Lugli GA, Duranti S, Turroni F, **Mancabelli L**, Ferrario C, Mangifesta M, Hevia A, Viappiani A, Scholz M, Arioli S, Sanchez B, Lane J, Ward DV, Hickey R, Mora D, Segata N, Margolles A, van Sinderen D, Ventura M. **"Bifidobacteria exhibit social behavior through carbohydrate resource sharing in the gut."** Sci Rep. 2015 Oct 28;5:15782. doi: 10.1038/srep15782.

27. Turroni F, Özcan E, Milani C, **Mancabelli L**, Viappiani A, van Sinderen D, Sela DA, Ventura M. **"Glycan cross-feeding activities between bifidobacteria under in vitro conditions."** Front Microbiol. 2015 Sep 24;6:1030. doi: 10.3389/fmicb.2015.01030. eCollection 2015.

28. Colagiorgi A, Turroni F, **Mancabelli L**, Serafini F, Secchi A, van Sinderen D, Ventura M. **"Insights into teichoic acid biosynthesis by *Bifidobacterium bifidum* PRL2010."** FEMS Microbiol Lett. 2015 Sep;362(17). pii: fnv141. doi: 10.1093/femsle/fnv141. Epub 2015 Aug 21.

29. Milani C, **Mancabelli L**, Lugli GA, Duranti S, Turroni F, Ferrario C, Mangifesta M, Viappiani A, Ferretti P, Gorfer V, Tett A, Segata N, van Sinderen D, Ventura M. **"Exploring**

Vertical Transmission of Bifidobacteria from Mother to Child." Appl Environ Microbiol. 2015 Oct 15;81(20):7078-87. doi: 10.1128/AEM.02037-15. Epub 2015 Jul 31.

30. Zanotti I, Turroni F, Piemontese A, **Mancabelli L**, Milani C, Viappiani A, Prevedini G, Sanchez B, Margolles A, Elviri L, Franco B, van Sinderen D, Ventura M. **"Evidence for cholesterol-lowering activity by *Bifidobacterium bifidum* PRL2010 through gut microbiota modulation."** Appl Microbiol Biotechnol. 2015 Aug;99(16):6813-29. doi: 10.1007/s00253-015-6564-7. Epub 2015 Apr 12.

31. Ferrario C, Milani C, **Mancabelli L**, Lugli GA, Turroni F, Duranti S, Mangifesta M, Viappiani A, Sinderen Dv, Ventura M. **"A genome-based identification approach for members of the genus *Bifidobacterium*."** FEMS Microbiol Ecol. 2015 Mar;91(3). pii: fiv009. doi: 10.1093/femsec/fiv009. Epub 2015 Jan 27.

32. Duranti S, Milani C, Lugli GA, Turroni F, **Mancabelli L**, Sanchez B, Ferrario C, Viappiani A, Mangifesta M, Mancino W, Gueimonde M, Margolles A, van Sinderen D, Ventura M. **"Insights from genomes of representatives of the human gut commensal *Bifidobacterium bifidum*."** Environ Microbiol. 2015 Jul;17(7):2515-31. doi: 10.1111/1462-2920.12743. Epub 2015 Feb 14.

33. Milani C, Lugli GA, Turroni F, **Mancabelli L**, Duranti S, Viappiani A, Mangifesta M, Segata N, van Sinderen D, Ventura M. **"Evaluation of bifidobacterial community composition in the human gut by means of a targeted amplicon sequencing (ITS) protocol."** FEMS Microbiol Ecol. 2014 Aug 13. doi: 10.1111/1574-6941.12410. Epub 2014 Sep 8.

34. Lugli GA, Milani C, Turroni F, Duranti S, Ferrario C, Viappiani A, **Mancabelli L**, Mangifesta M, Taminiau B, Delcenserie V, van Sinderen D, Ventura M. **"Investigation of the evolutionary development of the genus *Bifidobacterium* by comparative genomics."** Appl Environ Microbiol. 2014 Oct 15;80(20):6383-94. doi: 10.1128/AEM.02004-14. Epub 2014 Aug 8.

35. Milani C, Lugli GA, Duranti S, Turroni F, Bottacini F, Mangifesta M, Sanchez B, Viappiani A, **Mancabelli L**, Taminiau B, Delcenserie V, Barrangou R, Margolles A, van Sinderen D, Ventura M. **"Genomic encyclopaedia of type strains of the genus *Bifidobacterium*."** Appl Environ Microbiol. 2014 Oct 15;80(20):6290-302. doi: 10.1128/AEM.02308-14. Epub 2014 Aug 1.