Department of Biomedical Sciences, School of Medicine

Ph.D. Course in: Brain, Mind and Computer Science

Curriculum: Neuroscience, Technology, and Society

Series: XXXV

The Athlete's Microbiome Project: from molecular biology to bioinformatic

Coordinator: Prof. Anna Spagnolli

Supervisor: Prof. Antonio Paoli

Co-Supervisor: Prof. GianAntonio Susto

Ph.D. student : Laura Mancin

Thesis contents

PART 2: STUDY 1

PART 2: STUDY 2

I knew I wanted to be a scientist when I realized that science was not about the absolute truth.

Undertaking this doctorate program has been a truly life-changing experience.

I learned to always question, I learned to make mistakes but, above all, I learned that science and tiny organisms will be my lifelong partners.

I am extremely grateful to my supervisor Prof. Antonio Paoli for his vast wisdom and wealth of experience have inspired me throughout my studies and for the insightful conversations we had along the way.

I would thank Dr. Ian Rollo for his constant encouragement throughout my doctorate studies.

I emotionally dedicate this thesis to my family for the endless encouragement and support on my way to being a scientist.

Laura

Abstract

The gut microbiome represents a real "orchestra conductor" in the host's physio(patho)logy due to its implications in many aspects of health and disease. Reciprocally, gut microbiome composition and activity are influenced by many different factors, such as diet and physical activity.

The massive sequencing of gut microbiome specimens, thanks to the technological advancements in high-throughput sequencing and bioinformatics analyses, and more recently, the development of methods to quantify different microbial metabolites, allowed population-level studies to be carried out on the human microbiome. Their findings help to better understand the microbiome's role in physiology, its functional imbalance in various chronic pathologies [1] and, its implications on athlete's health and performance [2, 3]. In addition, mechanistic pre-clinical studies have brought novel insights into the underlying molecular mechanisms and can be used to test strategies, such as dietary challenges, to modulate the gut microbiota composition [4, 5].

The aim of our doctoral project was to first highlight the implications of gut microbiome on athlete's health and then, to focus the attention on the combination of physical activity and nutritional interventions to modulate gut microbiome composition in the context of health and performance (dietary strategies and food supplementation).

This new knowledge could be used to develop strategies (i.e., personalized dietary advice, supplementation, physical activity programs) to modulate the gut microbiome with the ultimate goal of improving performance in elite athletes and/or preventing various pathologies related to skeletal muscle metabolism.

The thesis is structured as follow:

1) **Study 1**.

   Type of study: Observational study

   Title: Athletes' Microbiome Project (AMP)

Issue: The characteristic of a 'healthy microbiome' remains unclear and population-based study reveal that interindividual variation is partially accounted by diet and lifestyle factors including exercise

Aim: Characterized the composition of gut microbiome in a well-defined population cohort (elite soccer players) and explore the associations of microbial features to metadata (dietary intakes).

2) **Study 2**.

Type of study: Randomized controlled trial

Title: The anti-inflammatory role of extra-dark chocolate on elite athletes' health and microbiome composition

Issue: High intensity and/or long-lasting exercise impairs the intestinal environment and the gut microbiome composition thus, consequently impacting health and performance

Aim: Identify potential food supplement (30g dark chocolate/day/1month) that may counteract the effect of intense exercise

3) **Study 3**:

Type of study Randomized controlled trial

Title: Effect of 30 days of ketogenic Mediterranean diet with phytoextracts on athletes' gut microbiome composition

Issue: Ketogenic diet seems to be detrimental for the athlete's gut microbiome

Aim: Investigating the effect of a specific ketogenic diet (KEMEPHY) on athlete's gut microbiome and health [6].

PART 1: BACKGROUND ON THE TOPIC OF DOCTORAL PROJECT

General Introduction to microbiome study

Gut Microbiome: What is it?

*In this thesis we will use both the term "microbiome" and "microbiota". Sometimes used interchangeably, these two terms have subtle differences. The microbi**ota** refers to the microbial taxa associated with humans to signify the communities of microorganisms within a specific environment, while microbi**ome**, on the other hand, refers to the collection of microbial taxa or microbes and their gene content.* Thus, if we consider these two terms differentially, "microbiota" is used to signify the communities of microorganisms, whereas "microbiome" is to signify the organisms and all their related genomes. Although we still need to confirm whether we are born sterilely [7], it is known that we begin to be colonized with microbes at birth, or our microbiota development begins well in amniotic fluid before delivery . Over the first several years of life, particularly during the first 3 years, our skin surface, oral cavity, and gut are colonized by a tremendous diversity of bacteria, archea, fungi, and viruses until the microbiota becomes adult-like [8]. The largest microbial community of the human microbiome is our intestinal tract harboring up to 100 trillion 10^{14} microbes, which are 10 times the number of human cells, and more than 150 times the number of human genes [8]. The vast majority of microbes reside in the colon with a density around 10^{11} to 10^{12} cells/ml. In 2016, Sander et al. [9] reestimated that the number of bacteria in adult colon is 3.8×10^{13} and the ratio of bacteria to human cells is closer 1:1 instead of 10:1 as previously estimated.

Microbiome research and DNA sequencing

The science of microbiome has a long history [10]. Historically, microbiology studies were almost entirely culture-dependent (it was necessary to grow the organism in the lab, outside their natural

habitat). However, it was a challenging process because many species were well adapted to live in human body and the suitable environment for human microbiome was not viable in *in vitro* condition. These challenges have resulted actually, in the past, in an underestimation of the complexity of the human microbial community. It was only in 2005, that the advances in DNA-sequencing technologies allowed researchers to analyze the DNA extracted directly from human sample, rather than from individually cultured microbes [11]. Moreover, in few years, DNA sequencing has fundamentally shifted away from classical Sanger automated sequencing to next-generation sequencing (NGS) analysis. Indeed, although the automated Sanger method is currently considered the "first-generation" sequencing technology, the sequenced reads produced were limited. For this reason, the shotgun sequencing techniques have been successively emerged to analyze longer fragments. Currently, Illumina /Solexa is the most widely used platform in the field of metagenomics.

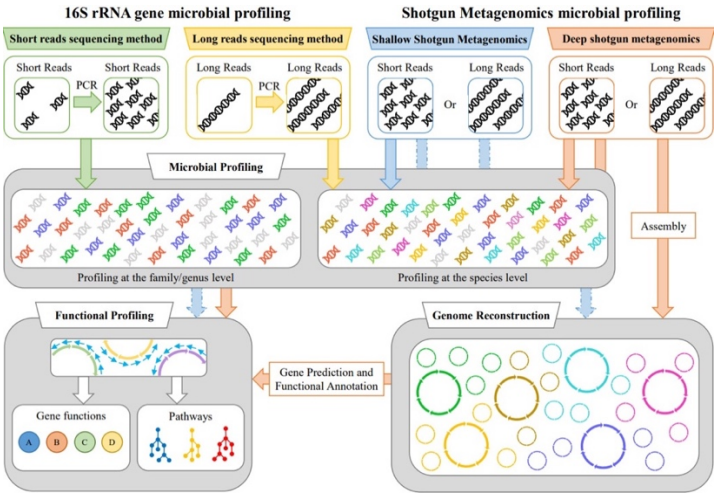
Introduction to Phylogenetics for bacteria analysis

Phylogenetics is the study of the evolutionary history and relationships among individuals or groups of organisms. Phylogenetics is important because it enriches our understanding of how genes, genomes, species and molecular sequences generally evolve. Human microbiome is very complicated with existing genetic and evolutionary relationships among species. This field of classification, identification and naming of biological organisms on the basis of shared characteristics is called taxonomy. Taxonomy stems from ancient Greek *taxis*, meaning "arrangement", and *nomia* , meaning "method". To understand the complexity of the human microbiome, it is important to recognize the genetic and evolutionary relationship between species. The Swedish botanist Carl Linneaus was known as the father of taxonomy, who developed a system for categorization of organisms, known as Linnaean taxonomy, and binomial nomenclature from naming organisms. Linneaus and others ranked all living organisms into seven biological groups or levels of classification: kingdom, phylum, class, order, family, genus, and species. There are no

domains in these classifications. The classification of domain is a relatively new grouping, which was first proposed by Woese et al. in 1997. They said, a formal or natural system of organisms should have a new taxon called “domain” above the level of kingdom. Archea, Bacteria and Eukarya are the three domains of life. At each lower level, organisms are classified with their most similar characteristics. Species-level analysis provides the most precise information of life; however higher-level analyses are also valuable, especially when species identification is challenging [12].

How to analyze microbial genome: 16S rRNA amplicon sequencing and Whole Genome Shotgun Sequencing (WGS)

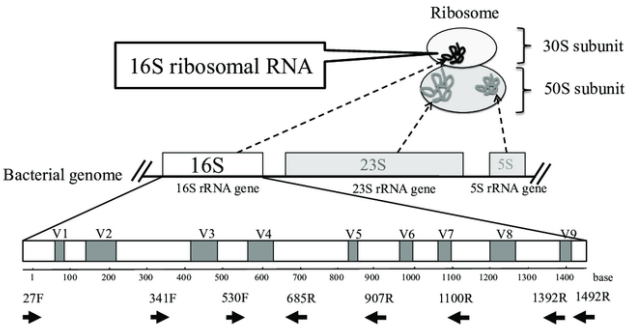
Currently, there are two main approaches to sequence the uncultured microbes: amplicon sequencing (in which one particular gene, often 16S rRNA, is amplified and sequenced) and random shotgun sequencing.



A breath of fresh air in microbiome science: shallow shotgun metagenomics for a reliable disentangling of microbial ecosystems, Microbiome Res Rep 2022;1:8.10.20517/mrr.2021.07

16S rRNA Sequencing

The 16S rRNA gene (or rDNA), is the conservative gene in the microbes.



Molecular Approaches to Studying Microbial Communities: Targeting the 16S Ribosomal RNA Gene, September 2016Journal of UOEH 38(3):223-232.DOI: 10.7888/juoeh.38.223

Lane et al. in 1985 [13] first described the use of 16S rRNA gene to identify and classify uncultured microbes in the environment. Because this gene has several desirable properties, the 16S rRNA gene sequencing has been the first metagenomics method. The properties include: 1) The 16S rRNA gene is ubiquitous and necessary component of ribosomes translating mRNA (it encodes the small subunit ribosomal RNA molecules of ribosomes that are used for the translation of mRNA into protein).; 2) The 16S rRNA gene contains highly conserved regions suitable for universal PCR primer design to amplify region of interests such as V1-V3 or V1-V4; 3) Well-studied primer sets are available for amplifying most organisms with high specificity for bacteria; 4) Well-curated databases of reference sequences and taxonomies are available allowing sequence comparison and taxonomic assignment of organisms; 5) The 16S rRNA gene sequencing in also relatively cheap and simple with mature analysis pipelines [14].

Limitation of 16S rRNA sequencing approach

1) Amplicon sequencing rRNA markers via PCR may miss detecting taxa due to various biases associated with PCR (which may result in substantially reducing microbial diversity in a community)

2) 16S rRNA sequencing overestimates the community diversity or species abundance due to the artificial sequence caused by sequencing errors and incorrectly assembled amplicons (i.e., chimeras)
3) Amplicon sequencing only discerns the taxonomic composition of microorganisms, but it cannot analyze the biological functions of associated taxa
4) Amplicon sequencing can only analyze the taxa that taxonomically informative genetic markers are known and can be amplified. It is difficult to be used for analyzing novel or highly divergent microbes
5) 16S rRNA sequencing approach lacks a golden standard for guiding decisions on quality control and statistical analysis methods (Xia and Sun 2017).

Shotgun Metagenomic Sequencing

Definition of Metagenomics

Shotgun metagenomic sequencing is a powerful alternative to 16S rRNA sequencing for analyzing complex microbiome communities and avoids some of these limitations. Metagenomics has been defined as “the genomic analysis of microorganisms by direct extraction and cloning of DNA from assemblage of microorganisms” and a metagenome has been defined as “the entire genetic information of an ensemble of organisms, living in a common habitat” [15]. In its approach, metagenomics refers to “computational methods that maximize understanding of the genetic composition and activities of communities” and metagenomic analyses have three basic tasks: taxonomy analysis (who are they?), functional analysis (what can they do?) and comparative analysis (how to compare them?).

Advantages of shotgun metagenomic sequencing

There has been an increasing interest in employing shotgun sequencing, rather than amplicon sequencing since this method has several advantages:

1) Shotgun metagenomics not only produces analysis data for generating hypothesis about the microbial community composition, but also provides a powerful tool to hypothesize microbial functions associated with different conditions, such as, health and disease, and treatment and control. Given the functional profile, researchers can generate hypotheses on community dynamics and metabolic properties [16]. In shotgun metagenomic sequencing, total DNA in a community is extracted and independently sequenced, which produces huge numbers of DNA reads that align to various genomic locations in the sample. Thus, after obtaining the shotgun metagenomic sequencing data, microbiome scientists can simultaneously explore two basic tasks of a microbiome study: which microorganisms are present within a sample and what each of them do. Indeed, shotgun metagenomics can fully characterize a community, including: i) community composition, ii) community member's genetic potential, iii) intra-species gene heterogeneity and iv) the metabolic potential of the community [17];
2) The shotgun metagenomic sequencing is potentially unbiased (accurate detection of all microorganisms and viruses [18]) so it has more chances to detect rare and novel species and viruses;
3) The shotgun metagenomic sequencing approach has the ability to discriminate strains of common species by gene content, which is not possible with 16S rRNA sequencing approach [16].

Bioinformatic methods in metagenomics

Two approaches are available: assembly-based and read-based metagenomics. In assembly-based metagenomics, separate reads are first *de novo* assembled to contigs and then clustered into so-

called genome bins during a binning process (see below). In read-based metagenomics, individual reads are classified with regard to taxonomy and function, thus it is suitable to analyze the taxonomical composition, functions of the metagenome, and metabolic pathways [19].

Processing of samples, DNA extraction and library preparation

Sampling

Sample is the first and most critical step in shotgun metagenomic studies. The first element of sample is to convert the source nucleic acid material into a sequencing library. Typically, there are several steps: first, to fragment long DNA into suitable size, then, to perform adapter addition. Last, to perform PCR to select molecules containing adapters at both ends and to generate enough quantities for sequencing. The most challenges of library preparation may be the quantitative biases and the loss of material occurred during preparation. However, to reduce the biases and loss of material, many algorithms have been developed and steps have been taken.

DNA extraction

DNA extraction in shotgun metagenomic sequencing method relies on the extraction of DNA from all cells in a sample and the obtained amounts of high-quality nucleic acids must be sufficient. The resulted DNA sequences reads are aligned to various genomic locations for the myriad genomes present in the sample.

Assembly

The step of assembling allows to assemble short reads into longer, contiguous sequences (‘contigs’) which permit downstream bioinformatic analysis smoothly. There are two types of assembly: reference-based assembly and de novo assembly. The first one performs well if the closely related reference genome sequences are available in the metagenomic datasets, while it performs poorly if

in the sample genome exists a large insertion, deletion, or polymorphisms. The novo assembly typically requires larger computational resources (more memory and run times).

Binning

Binning is defined as the process of sorting DNA sequences into groups that might represent an individual genome or genomes from closely related organisms. Mixtures of variable length of sequence fragments originating from various organisms returned by contig assembly. It is a challenge for assembly to reconstruct entire genomes, thus, following by assembly, it is necessary to bin genome fragments.

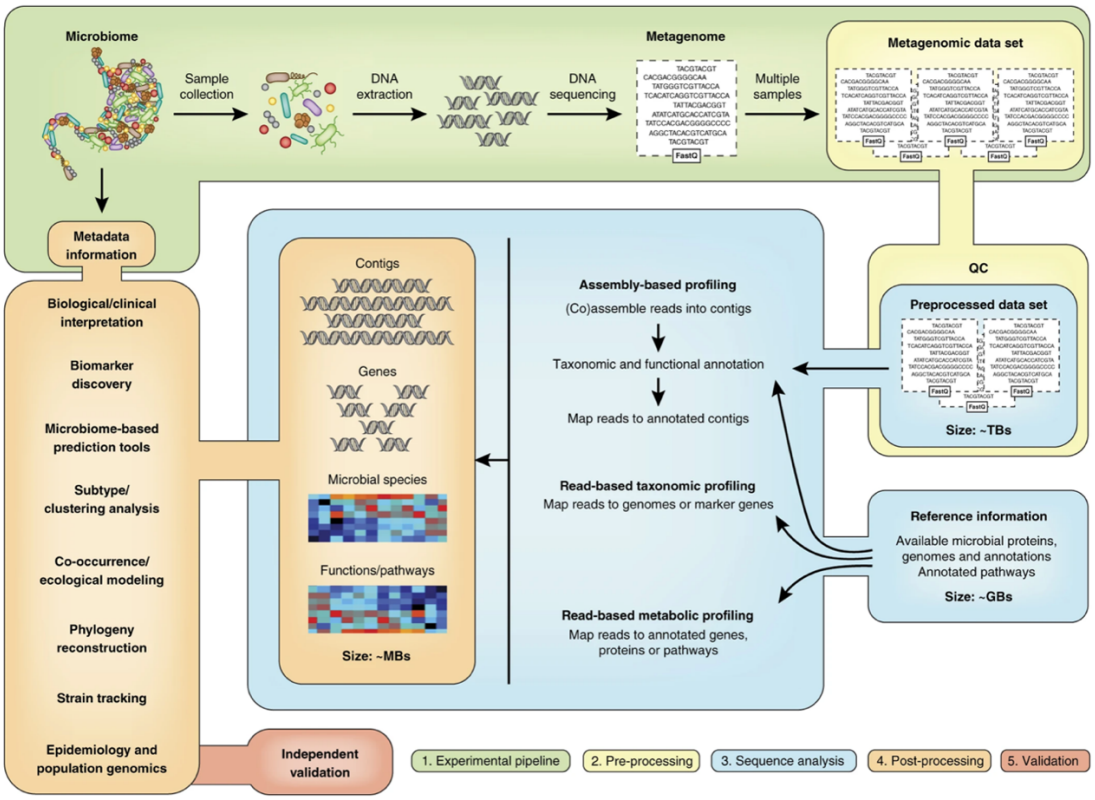
Annotation

To gain insight beyond taxonomic composition, the sequenced need to be annotated. Metagenome functional annotation of metagenomic sequences generally has two non-mutually exclusive steps: the gene prediction and annotation. Gene prediction refers to the procedure of identifying gene of interest, protein and RNA sequences coded on the sample DNA; functional annotation of metagenomic datasets assign putative gene functions and taxonomic neighbors.

Limitation (challenges) of shotgun metagenomic data

Despite the advantages, compared to 16S rRNA amplicon sequencing, the bioinformatic analysis of shotgun metagenomic data has some challenges. As an example, the data sets generated are very large and thus, highly complex to analyze (this makes, obviously, the bioinformatic analysis very complicated). Further, the cost of whole-genome sequencing is still high, compared to 16S (approximately 160$ for shotgun metagenomics analysis and 50$ for 16S rRNA amplicon analysis).

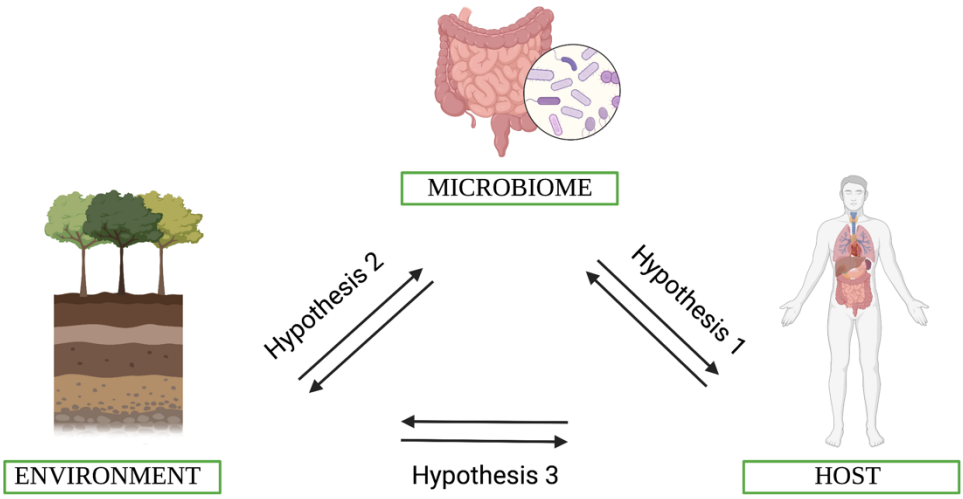
Schematic summary of shotgun bioinformatic workflow:



Quince, C., Walker, A., Simpson, J. et al. Shotgun metagenomics, from sampling to analysis. Nat Biotechnol 35, 833–844 (2017). https://doi.org/10.1038/nbt.3935

Research themes and statistical hypothesis in human microbiome studies

There are two themes in the current microbiome studies: 1) to characterize the relationship between microbiome features and biological, clinical, or experimental condition; and 2) to identify potential biological and environmental factors that are associated with microbiome composition. The goal of these studies is to understand mechanisms of host genetic and environmental factors that shape microbiome. Insights gained from studies potentially contribute to the development of therapeutic strategies in modulating the microbiome composition in human diseases [20]. The interaction among environment, microbiome and host are dynamic and complicated. In the following figure we schematically represent the interactions.



To study interactions among environment, the general research hypotheses could be developed. Hypothesis 1 is to test the association between microbiome and host: whether the composition of the microbiome or "dysbiotic" microbiome is linked to the health or disease of host, i.e., the association between microbiome and host physiology. For example, in inflammatory intestinal bowel disease (IBD), it is hypothesized that dysbiosis is associated with progression and symptoms of the disease. Hypothesis 2 is to test whether microbiome is associated with environmental or biological covariates, whether environmental factors impact microbiome, or whether an intervention influences a specific microbiome composition (diversity) in health and disease. For example, it may be assessed whether dietary interventions shape gut microbiota or whether probiotic intervention impacts the composition of the human microbiota. Hypothesis 3 is to investigate whether environmental features may be associated with host. In general, the statistical null hypothesis could be: there is no difference of microbiome composition in different experimental groups (i.e., health and disease). The core theme of these statistical hypothesis could be the same, i.e., to explore the impacts of external factors (interventions) on microbiome composition and/or richness of microbiota. However, these research topics are varying among alpha diversity (species diversity in each individual sample), bacterial richness, phylogenetic diversity, and species evenness in each sample [21].

Insights into gut microbiome, diet, and athlete's physiology

As it previously discussed, a growing plethora of studies show how specific microbiome signatures are associated with lifestyle, diet, and many diseases states such as inflammatory bowel diseases (IBDs), cancer, alcoholic and non-alcoholic fatty liver diseases, type 2 diabetes, obesity, and mental related conditions. However, few studies reveal association between exercise and specific microbiota profiles with the majority lacking causation or mechanistic explanation. It seems that communication between gut microbiota and athlete's physiology is still "hidden" and we do not have interpretable strategies able to decipher the secret information. However, thanks to new meta-omics analysis and novel computational tools, the identification of this bidirectional influence is becoming clearer. Machine learning might offer an opportunity to reveal how bacteria communicate with host exercise physiology, highlighting the potential relevance of our microbiota in sport science [2]. The gut microbiome density varies along the gastrointestinal (GI) tract. It is low in the stomach, duodenum and jejunum and increases in the ileum and colon. The stomach is characterized by the presence of oxygen and high acidity. In the small intestine, the microbiota is mainly composed of facultative anaerobic bacteria (e.g., Lactobacilli, Streptococci and Enterobacteria) and strict anaerobic bacteria (Bifidobacteria, Bacteroides and Clostridia). In the colon, the slower intestinal peristalsis and anaerobiosis favor the onset of a large ecostystem with the highest bacterial diversity and density, where Firmicutes and Bacteroidetes account for 80–90%. The gut microbiota has many beneficial functions in the organism when it is in symbiosis with the host. For instance, the bacteria play a pivotal role in digestion by ensuring the fermentation of substrates and non-digestible food residues (such as fiber), by facilitating the nutrient assimilation thanks to a set of enzymes that are not present in human cells, and by participating in the synthesis of some vitamins. They also influence the overall functioning of the GI tract and participate in the functioning of the intestinal immune system, which is essential for the intestinal wall barrier function. In conclusion, the human gut microbiota is a complex ecosystem that is different from that

of other microbiota types in the human body. Due to the complexity of this ecosystem, host-microbiota interactions/axis represent a fragile equilibrium that can be disrupted in many factors and pathologies [22].

Key points:

- The trillions of microbes in the gastrointestinal tract play essential roles both in health and disease.
- Different dietary strategies and exercise training methods may alter the composition and functional activity of the intestinal ecosystem.
- We cannot define one healthy microbiota scenario for all athletes, however, there are several opportunities to modulate gut microbiota and consequently athlete's health.
- To get a better understanding of how microbiota communicate with athlete's physiology it is required to go beyond correlation studies towards the combination of meta omics technologies with computational tools.
- We explore the possibility to leverage Machine Learning models to predict exercise-induced alteration and performance measures (e.g. VO_{2peak} and VO_{2max}, glycemic and markers of inflammation), to automatically distinguish between responders and non-responders to specific nutrients or supplements (i.e., protein digestion), and to get additional information about which are the most relevant features and correlations to solve the mentioned tasks.

Diet influences the gut microbiota composition

See "***Mancin L****, Rollo I, Mota JF, Piccini F, Carletti M, Susto GA, Valle G, Paoli A. Optimizing Microbiota Profiles for Athletes. Exerc Sport Sci Rev. 2021 Jan;49(1):42-49. doi: 10.1249/JES.0000000000000236. PMID: 33044333*" for further information.

General dietary guidelines for athletes are described in consensus statements from different scientific societies; moreover, each athlete has a specific dietary requirement based on the demands of exercise and individual training/performance goals.

Indeed, training can be broadly divided into endurance or strength/power based and,

it is intuitive that different nutritional approaches are necessary to support the adaptation to the different training stimulus. To this end, different nutritional strategies are adopted to optimize the specific athletes' energy requirements, body composition targets and performance goals.

Carbohydrates:

The effects of carbohydrates on gut microbial communities vary from the type and the amount of these macronutrients. While diets high in fermentable carbohydrate improve the growth of short chain fatty acids (SCFAs) - producing bacteria (i.e.: Prevotella species), instead, dietary patterns low in microbiota-accessible carbohydrates (MACs) promote the growth of mucus- degrading bacteria, impairing the gut intestinal barrier. However, although athletes may gain benefit from high intakes of MACs, some individuals may experience undesirable side effects such as bloating and flatulence, thus impacting exercise performance. Since the tolerance and the response efficiency to fiber is person-specific and correlates with the presence of certain fiber-degrading species, it would be interesting to characterize the microbiota composition of athletes, categorizing them as "responders" and "non-responders" to fiber intakes. Then, once clustered, it may be reasonable to target specific type and amount of fiber able to positively modulate microbiota without raising side effects.

Protein:

To date, the evidence concerning the effect of protein on gut microbiota composition is still contradictory. A high intake of protein increases the abundance of proteolytic microbes and derived end-products that can negatively impact barrier function and host physiological response. However,

microbial proteolytic fermentation occurs in specific situations, such as when the intake of protein is elevated and when low fiber - low resistant starches are provided. Interestingly, the MACs found within specific foods have the ability to suppress protein fermentation by lowering pH in distal gut and decreasing the requirement of amino acids as energy source for microorganisms. What deserves particular attention is the quality of protein sources: dietary patterns based on red meat, contain higher levels of choline, phosphatidylcholine and L- carnitine, which are bioactive precursors converted by gut bacteria in trimethylamine (TMA), the precursor of trimethylamine N-oxide (TMAO), an organic compound associated with the risk of atherothrombotic cardiovascular disease (CVD). While the mechanisms of how TMAO can develop CVD are still unknown, several studies are testing whether CVD may be controlled by reducing the relative abundance of TMAO-producing bacteria . Thus, for those athletes who consume high intake of red meat or have a family story of CVD, it may be beneficial to investigate the presence of TMAO-producing bacteria to eventually suggest dietary strategies low in TMA 's precursors. However, it has to be underlined that athletes usually reach the higher required protein intake through the use of dietary protein supplementation derived from milk proteins.

Fats:

See: *Paoli A, **Mancin L**, Bianco A, Thomas E, Mota JF, Piccini F. Ketogenic Diet and Microbiota: Friends or Enemies? Genes (Basel). 2019 Jul 15;10(7):534. doi: 10.3390/genes10070534. PMID: 31311141; PMCID: PMC6678592* for further information.

Consumption of a high-fat diet (HFD) significantly reduces the fecal concentration of short-chain fatty acids (SCFA), including butyrate, and of Bifidobacteria, compared with a low-fat diet. Moreover, several human studies demonstrated that HFDs increase the total anaerobic microflora and Bacteroides. However, by definition, in HFDs, the carbohydrate amount in the total energy intake is decreased. Therefore, it is not clear whether microbiota composition and metabolism are

mainly influenced by elevated fat or reduced carbohydrate (in the form of fiber content). In addition, more than their amount, the fat quality plays an important role in the gut microbiota composition. For example, we demonstrated that 30 days of ketogenic Mediterranean diet, mainly composed of polyunsaturated fatty acids (PUFAs), and especially of n-3 PUFAs modulate the intestinal microbiota in a beneficial way.

Probiotic

*See: "*de Paiva AKF, de Oliveira EP, **Mancin L**, Paoli A, Mota JF. Effects of probiotic supplementation on performance of resistance and aerobic exercises: a systematic review. Nutr Rev. 2022 Aug 11:nuac046. doi: 10.1093/nutrit/nuac046. Epub ahead of print. PMID: 35950956*."* for further information.

Probiotics are live microorganisms that, when administered in adequate amounts, confer a health benefit to the host. Their effects on the gut and immune system are the most researched applications. For example, fermented foods containing lactic acid bacteria, such as milk products and yogurt, represent a source of ingestible microorganisms that may beneficially regulate intestinal health and even treat or prevent IBDs. Lactobacillus, Bifidobacterium, and Saccharomyces strains have been safely and effectively used as probiotics for a long time. Roseburia spp., Akkermansia spp., Propionibacterium spp. and Faecalibacterium spp. are also promising probiotic microorganisms. In line with these findings, we conducted a systematic review to assess the effects of probiotic supplementation on athletic performance.

Bioactive Non-Nutrient Plant Compound

See*: "Sorrenti V, Ali S,* ***Mancin L****, Davinelli S, Paoli A, Scapagnini G. Cocoa Polyphenols and Gut Microbiota Interplay: Bioavailability, Prebiotic Effect, and Impact on Human Health. Nutrients. 2020 Jun 27;12(7):1908. doi: 10.3390/nu12071908. PMID: 32605083; PMCID: PMC7400387."* for further information.

Some bioactive non-nutrient compounds present in fruits, vegetables, grains, and other plants have been linked to a reduction in the risk of major chronic diseases. These plant compounds include prebiotics and probiotics, as well as several chemical compounds, such as polyphenols (the largest group) and derivatives, carotenoids, and thiosulfates. Polyphenols can be subclassified into four main groups: flavonoids (including eight subgroups), phenolic acids (e.g., curcumin), stilbenoids (e.g., resveratrol), and lignans. They promote health by limiting oxidative stress. Common polyphenol-rich food types include fruits, seeds, vegetables, tea, cocoa products, and wine. The relative abundance of Bacteroides is increased in people consuming red wine polyphenols. Moreover, it has been reported that the abundance of pathogenic Clostridium species (C. perfringens and C. histolyticum) is reduced after regular consumption of fruit, seed, wine, and tea polyphenols. In a recent review, Rajha et al. (2021) showed that polyphenol metabolites interact with gut microbiota and mitochondria to fight many diseases, such as obesity, depression, inflammation, and allergy. According to this, we performed a systematic review investigating the potential effect of cocoa flavonoid on gut microbiota and the associated host health benefits.

Vitamins

Some vitamins are directly produced by the gut microbiota, and others play a role in modulating the presence of beneficial/detrimental bacterial species. Specifically, vitamin A can modulate health-beneficial microbes of the Bifidobacterium, Lactobacillus and Akkermansia genera. Some B-complex vitamins are produced by gut commensals, and some of them contribute to increasing the virulence/colonization of potentially pathogenic microbes. Vitamin C, D, and E supplementation may alter the microbiota composition by increasing the concentration of beneficial species, such as Bifidobacterium and Lactobacillus. Thus, vitamin intake could have a significant role in modulating gut microbiota. Moreover, this effect might depend on the host's pre-supplementation vitamin level. However, clinical trials are still necessary to avoid adverse effects due to excess vitamin intake.

Potential Links between Gut Microbiome and Physical Fitness/Sports Performance

See "***Mancin L**, Rollo I, Mota JF, Piccini F, Carletti M, Susto GA, Valle G, Paoli A. Optimizing Microbiota Profiles for Athletes. Exerc Sport Sci Rev. 2021 Jan;49(1):42-49. doi: 10.1249/JES.0000000000000236. PMID: 33044333*" for further information.

The Athletes' Gut Microbiota, a Specialized Microbiota?

It is acknowledged that gut microbiota changes depend on individual factors, particularly in athletes, including energy expenditure, diet and drug intake (especially antibiotics or Non-Steroidal Anti-Inflammatory Drugs). A great body of evidence shows that the gut microbiota of athletes is different from that of other populations and displays higher microbial diversity. In 2014, Clarke et al. were the first to demonstrate that microbial diversity is increased in elite rugby players compared with matched controls. Specifically, the abundance of the phylum Bacteriodetes was decreased, whereas that of the genus *Akkermensia* was increased in athletes with low body mass index (BMI) (<25 kg/m^2) compared with the high BMI (>28 kg/m^2) group. Moreover, Estaki et al. showed that peak oxygen uptake

(VO_2peak), the gold standard measure of cardiorespiratory fitness, can account for more than 20% of the variation in taxonomic richness in healthy men and women, after adjusting for all other factors, including diet. Indeed, the abundance of key butyrate-producing taxa (*Clostridiales*, *Roseburia*, *Lachnospiraceae*, and *Erysipelotrichaceae*) was increased in individuals with high VO_2peak values. In 2019, Scheiman et al. showed that the relative abundance of *Veillonella* was increased after a marathon and that the inoculation of a strain of *Veillonella atypica* from runner stool samples into mice significantly increased exhaustive treadmill run time in the inoculated animals. They also demonstrated a mechanistic link of the change in performance with lactate metabolism. The recent reviews by Mohr et al. in 2020 and by Aya et al. in 2021 concluded that in most cases, the alfa and beta diversity of gut microbiota are not different among sports disciplines, but some differences can be highlighted for some genera or taxa abundances. For example, O'Donovan et al. concluded that microbial diversity does not differ among sport disciplines but, on the other hand, they observed a greater abundance of *Bifidobacterium animalis*, *Lactobacillus acidophilus*, *Prevotella intermedia* and *F. prausnitzii* in athletes with high dynamic components (high VO_2max), and greater abundance of *Bacteroides caccae* in athletes with both high dynamic and static components (in relation with the maximal voluntary contraction component). Besides their chronic training regimes, the dietary intake patterns of athletes are often different from those of sedentary subjects, as is medication intake. These factors also might influence their gut microbiota composition. Finally, some data show that prolonged excessive exercise could have a detrimental effect on intestinal function. Indeed, strenuous, and prolonged exercise increases intestinal permeability and alters the gut barrier function. This promotes bacterial translocation from the colon, leak of bacterial LPS into the bloodstream, and activation of systemic inflammation. GI symptoms (e.g., abdominal pain, nausea, and diarrhea) are reported by 70% of athletes after strenuous exercise, and the frequency is higher in elite athletes than in recreational exercisers. Besides the overall "healthy" gut microbiota in athletes, many discrepancies can be observed in the microbiota profiles at lower taxonomic levels in relation to many confounding factors linked to the exercise type (e.g.,

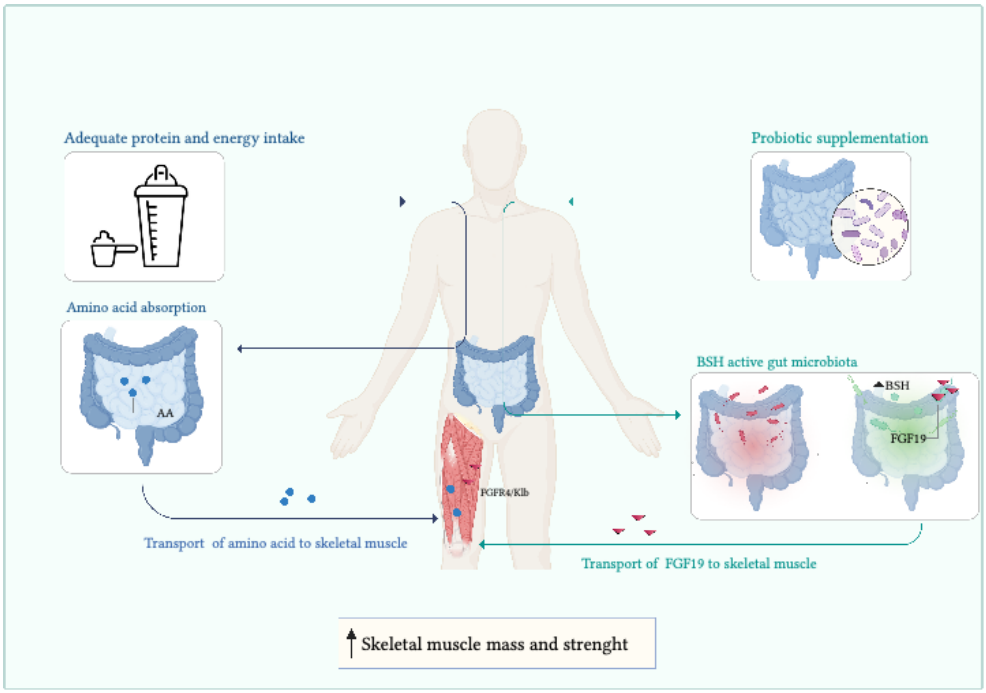
intensity, mode, contraction type, duration, frequency), diet, drug intake, environment, season, sleep and many others.

The Gut- Skeletal Muscle Axis

See "***Mancin L**, Wu GD, Paoli A. Gut microbiota-bile acid-skeletal muscle axis. Trends Microbiol. 2022 Oct 29:S0966-842X(22)00286-4. doi: 10.1016/j.tim.2022.10.003. Epub ahead of print. PMID: 36319506*." for further information.

The gut microbiota represents a 'metabolic organ' that can regulate human metabolism. Intact gut microbiota contributes to host homeostasis, whereas compositional perturbations, termed dysbiosis, are associated with a wide range of diseases. Recent evidence demonstrates that dysbiosis, and the accompanying loss of microbiota-derived metabolites, results in a substantial alteration of skeletal muscle metabolism. As an example, bile acids, produced in the liver and further metabolized by intestinal microbiota, are of considerable interest since they regulate several host metabolic pathways by activating nuclear receptors, including the farnesoid X receptor (FXR). Indeed, alteration of gut microbiota may lead to skeletal muscle atrophy via a bile acid–FXR pathway. The recent review by Mancin et al., aims to suggest new strategies aimed at optimizing skeletal muscle functionality. Figure 1.

Figure 1. Probiotic and adequate protein intake for the maintenance of skeletal muscle health.



On one side, adequate daily energy-protein intake improves musculoskeletal health. Protein digestion and amino acid absorption represent the sequential mechanism by which ingested protein provide amino acid available for the organism. Once ingested, amino acids are released and taken up across the gut mucosa, where many of them, are released into the systemic circulation to be transported and taken up by skeletal muscle. On the other side, probiotics may act synergically. Specific probiotics positively modulate gut microbiota (by preventing and/or treating gut dysbiosis) whole promoting the growth of beneficial bacteria (i.e.: increasing the relative abundance of BSH-containing bacteria). Beneficial bacteria consequently metabolize conjugated bile acids, which can up-regulate the FXR-FGF15/19 signaling. FGF15/19, once released into the systemic circulation, can be transported, and taken up by skeletal muscle which expresses the receptors for FGF15/19 (FGFR4/BetaKlotho). FGF15/19 increases muscle mass, myofiber size and muscle strength. Abbreviation can be found in the text.

The future meta-omics potential

The recent "omics" technological advances led to the understanding that the human responses to exercise may be driven by external factors such as diet, lifestyle, environment and microbiota composition. Since computational system approaches have the potential to lead to medical and technological breakthroughs, they may also provide a great opportunity for sport science to understand the functional impact of microbiota on athlete's physiology. By implementing different analyses, from metagenomics to meta-transcriptomics, meta- proteomics and meta-metabolomics, we can effectively address the complexity of the microbial influence on host physiology and exercise. However, although some efforts are being made, the acquisition of new biomedical insights from the combination of these omics platforms is still hampered by their high cost and data heterogeneity. To obviate this issue, it may be convenient to identify which specific meta-omics analysis could be adopted and combined with computational strategies aiming to understand which microbes are doing what. In sport science, since the purpose is understanding the systemic mechanisms of gut microbiota's influence on athlete's physiology, it may be pioneering to combine metagenomics and meta-metabolomics analysis with innovative mechanistic models able to provide access to this "unknown" mechanism.

Microbiota and personalized nutrition (a machine learning-based approach)

Machine learning for microbiome science: a utopia?

Machine learning (ML) approaches represent powerful tools in data-intensive application because they do not require a prior understanding of the underlying mechanisms governing the physical phenomenon under examination, but they automatically learn from the collected data. To date, ML tools have been broadly used in biomedical research and life science, in order diagnose or predict the risk of cancer, cardiovascular diseases, genetic disorders and, more generally to produce health

outcome predictions.

More specifically for sport science, ML approaches can be adopted to support and enhance researches investigating the connection between the microbiome and exercise. Despite the application of microbiome analysis in sport sciences, ML is still in its infancy (and, thus, calls for a "healthy dose of skepticism"), it can provide profound insights on how athlete's physiology is influenced by several different factors: ML may play a key role for such purpose.

For instance, a ML model could predict the athlete's exercise responsiveness in terms of glucose homeostasis and insulin sensitivity or in terms of biomarkers signature of aerobic fitness (ie., VO_{2peak}); such prediction may be used to provide customized lifestyle recommendation for modulating an individual's microbiota and consequently improving the athlete's responsiveness to exercise and the general health. For this purpose, the interpretability of the adopted ML model will be crucial as it will enable the user to get additional knowledge, besides predictions, about which are the most relevant features to solve the task at hand. Under the assumption that the ML model has satisfactory performance in terms of predictive power, the factors deemed as the most important by the model will be the ones on which the athlete should focus more, in order to maximize his responsiveness to exercise and improve his health. Despite the appeal of ML in this context, it must be stressed that the design and deployment of a ML-based solutions is far from being trivial. In general, any application domain where ML approaches are not yet fully established should undergo a preliminary phase devoted to a correct and grounded formalization of the problem. This calls for interdisciplinary collaborations in order to establish a common ground aimed at facilitating knowledge transfer among researchers from different fields. From a more technical perspective, a number of aspects must be considered in order to avoid any misuse of the proposed solutions and misinterpretation of the obtained results. Some important considerations include (but are not limited to) the following points:

- The data collection process should be carefully designed in order to ensure the storing of easily measurable and informative quantities. Particular attention should be directed to the

above mentioned high inter-variability of microbiome features and, specifically, data used to train the ML model should be fully representative of the population being analyzed.

● The target variable to be predicted should be a quantitative measure and as much as possible an objective one. Qualitative target variables should be replaced with adequate quantitative proxy variables and problems where subjectivity might affect some of the recorded measurements should be tackled by exploiting ML models that are robust to such sources of uncertainty (i.e.: VO_{2peak} and VO_{2max}, glucose and lactate blood level).

● As mentioned before preference should be given to interpretable ML models, i.e. models whose inherent logic is easily understandable by human beings. This poses a fundamental challenge since the best performing models (such as Deep Learning models) are often the hardest to be explained, but a growing number of researchers in the ML community are currently focused on the development of interpretability methods for so-called "black-box" models. This will enable a deeper comprehension of the problem at hand by identifying, for instance, which are the features (that can be represented by metagenomics, metabolomics, clinical and non-genetic information) that have the highest impact on the final prediction produced by the model. As an example, Zeevi and colleagues (2015) found that the gut microbiota has a great influence on the postprandial glycemic response of an 800-person cohort. The authors, using personal data and microbiome profiles, devised a ML algorithm which, integrating blood parameters, dietary habits, anthropometrics, physical activity, and gut microbiota features, accurately predicts personalized postprandial glycemic response to real-life meals. Similarly, Liu and colleagues (2020) revealed that baseline gut microbiota composition can accurately predict personalized exercise response in subjects with prediabetes. Both the abovementioned studies exploited ML models combined with interpretability methods to get feature importance measures, but focused only on how each feature affects the model output individually (i.e. independently from the others). We

advocate the need to extend such analysis and take into account also the impact that pairs (or groups) of features have on the produced predictions.

It should be underlined that the use of intelligent algorithms should always be intended as a tool in support of decision-makers; the concept of personalized sport nutrition does not have to replace the role of physician, dietician or nutritionist and must be pertinent to the physical, clinical and emotional needs of athletes. In the context of personalized sport nutrition, ML-based solutions should be seen as a powerful tool for medical staff which holds the final approval and takes part in looking at compliance.

An integrative approach: Microbiota and personalized nutrition

The challenge of personalized sport nutrition lies in the ability to systematically prove that the modulation of specific features of the athlete's diet (microbiome) is responsible for improving health and performance. To achieve the goal of personalized nutrition, these tailored nutritional recommendations might be firstly suggested for the amelioration of individual's gut health. Claimed that i) similar foods have different effects on different individuals' microbiome ; ii) gut microbiome changes are related with the personal baseline microbiota profile; iii) diet influences host response in an individualized way that may be predicted by the personal microbiome composition iv) intrinsic and extrinsic factors affect microbiota composition; personalized prediction might be developed through ML tools in favor of ameliorating the athletes' response to a specific nutrient, supplement, or medication. In athletes, it could be a great challenge to predict the response to prebiotics, probiotics, symbiotic supplementation, considered that the colonization of probiotic is strictly associated and predictable by "pre-treatment microbiome profile". Thus, it may be useful to identify person-specific microbiota configuration to suggest personalized therapies enabling persistent live-bacteria colonization.

In addition, since the microbiome plays a huge role in protein digestion, amino acid absorption and,

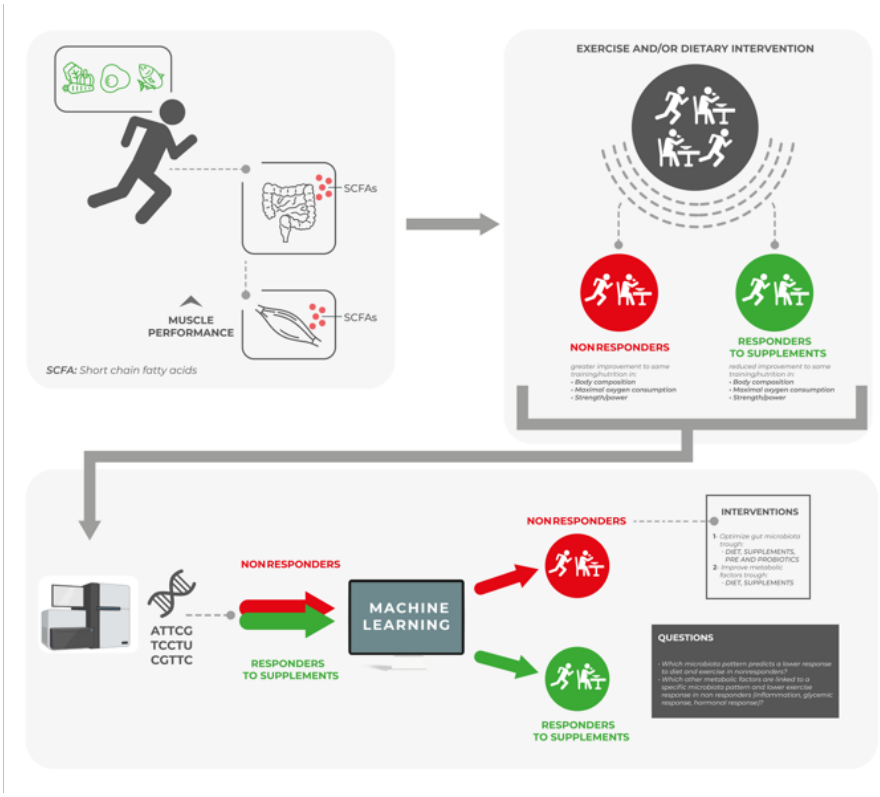
as such, on the bioavailability of the ingested protein derived amino acids, it may be interesting to discover the personal-specific tolerance and diversification in the physiological effect of whey protein or casein in order to personalize as much as possible the individual protein supplementation. For this purpose, ML models trained to distinguish between responders and non-responders (through the formalization of a binary classification task) may

help in achieving better results, thanks to their ability to process multivariate data effectively. In sports science, personalized fluid and electrolyte recommendations has been achieved by exploiting measures of sweat losses during exercise (Gatorade Gx system).

Harnessing these emerging technologies might represent an exciting promise for improving public health by optimizing diet and personal lifestyle. Whereas, in sports these technologies

may be utilized to maintain or improve performance. Figure 2.

Figure 2: Workflow for potential meta ‘omics studies in athletes



Mancin L, Rollo I, Mota JF, Piccini F, Carletti M, Susto GA, Valle G, Paoli A. Optimizing Microbiota Profiles for Athletes. Exerc Sport Sci Rev. 2021 Jan;49(1):42-49. Doi: 10.1249/JES.0000000000000236. PMID: 33044333.

Conclusion and perspectives

Today, there is no doubt that the discovery of the gut microbiota community opened a promising and rapidly growing research field on the potential beneficial health effects of manipulating the gut microbiota. Indeed, the gut microbiota influences the function of the intestine and also brain and metabolic tissues, such as skeletal muscle. Although it is established that the microbiome may influence the development and treatment of several human diseases, less literature is available about the specific functional contribution of microbiota to an athletes' physiology. Due to the complexity and variability in sports' tasks (and related specific nutritional approaches), and the greater inter-variability of microbiome features (and related response to diet) it is not reasonable to speculate that a unique "microbiota healthy configuration" could be defined for different sports and athletes. Indeed, to date, there are many limitations in the current research that so far need to be considered. It is by now intuitive to recognize the necessity to conduct long-term studies on athletes (at different time-point and scenarios), to use appropriate analyses such as metagenomics coupled with metabolomics, and to adopt innovative ML strategies capable of investigating the functional mechanism of the two-way crosstalk of microbiome and exercise. This will be made possible thanks to the growing research interest, within the ML community, in the design of methods to make black-box models more interpretable. Considering that there is no one optimal microbiota composition for athletes, we predict that meta omics data combined with interpretable machine-learning approaches will reveal how microorganisms interact with each other and with their host in order to identify different healthy microbiota scenarios relevant for athlete's health and performance. In many countries, National Gut Human Projects (e.g., the American Human Microbiome Project and the European Human Microbiome Action that started in 2021) have been set up to collect human fecal samples and to correlate the obtained microbiota results with the host's characteristics. Human fecal samples and metagenomic data are currently collected for future biostatistics analyses. For instance, the Million Microbiome of Humans Project (MMHP) is a major international project, the aims of which are to create the largest human microbiota database in the

world, to analyze 1 million samples, and to explore the full microbiome potential. One major milestone will be to launch observation studies with well-defined cohort of athletes as well as interventional studies to modulate the gut microbiota composition because, as we demonstrated, the gut can adapt its bacterial community in response to external factors, such as nutrition and physical activity.

The scientific and medical communities must now find the best way(s) to optimize the nutrition-gut microbiota-physical activity triad for each patient or athlete. The possibilities offered by 'biotic nutrition' represent a veritable "playground" for scientists. The challenge is to develop innovative, original and promising microbiota-based strategies to optimize sports performance and medical treatments or to delay disease onset. Finally, it is essential to increase the population's awareness of the need for a healthy diet and some physical activity for a healthy microbiota, although the triad mechanisms have not been fully elucidated yet. Some scientific organizations and large food companies are already campaigning about the importance of a healthy diet as a key factor in microbiota formation. However, they did not include the physical activity component. Indeed, the goal should be to make clear that both a healthy/well-balanced diet and regular (high-level) physical activity practice are needed to improve gut microbiota composition/function for better health and/or performance. We think that outreach programs should also include the triad concept to develop individualized microbiota-based strategies for health and sports performance management.

PART 2: EXPERIMENTAL DATA OF DOCTORAL PROJECT

The studies included in this section, all apply bioinformatic methods in metagenomics.

The aim of our research studies, one observational and two randomized controlled trials, was to investigate the implication of gut microbiome on athlete's health and then, to develop strategies (i.e., personalized dietary suggestions), to modulate gut microbiome, and, in turn, improving athlete's performance.

STUDY 1: Athlete's Microbiome Project (AMP)

See our systematic reviews for the general background.

1) **Mancin L**, Rollo I, Mota JF, Piccini F, Carletti M, Susto GA, Valle G, Paoli A. Optimizing Microbiota Profiles for Athletes. **Exerc Sport Sci Rev**. **2021** Jan;49(1):42-49. doi: 10.1249/JES.0000000000000236. PMID: 33044333.

2) **Mancin L**, Wu GD, Paoli A. Gut microbiota-bile acid-skeletal muscle axis. **Trends Microbiol**. **2022** Oct 29:S0966-842X(22)00286-4. doi: 10.1016/j.tim.2022.10.003. Epub ahead of print. PMID: 36319506.

Introduction to AMP:

Dietary contributions to health and chronic conditions, such as obesity, metabolic syndrome, cardiovascular diseases, and age-associated sarcopenia are of universal importance. For example, age-associated sarcopenia and related mortality/morbidity have risen dramatically over the past decades [23], with the gut microbiome implicated as one of several causal human-environment interactions. According to the great importance of microbiome on skeletal muscle metabolism and athlete's health, its role has been recently considered in the field of sport and exercise. Recent scientific advances suggest that nutrition influences athletic performance via the gut and the trillions

of microorganisms that inhabit this ecosystem. Importantly, diet affects the microbial community within the gut and, as a result, the gut microbiota mediates many of the effects of diet and nutrition on health. Indeed, given the microbiota's potential to influence athletic health and performance, "fueling the microbes" should be seen as a pivotal strategy for athletes attempting to optimize performance. Surprisingly, the details of the microbiome's role in athletes' health and performance have proven difficult to define reproducibility in large athlete populations, probably due to the complexity of dietary patterns, the effects of individual characteristics, the difficulty to perform microbiota analysis in elite sport setting and the personalized nature of the gut microbiome. To overcome these challenges, we launched the Athlete Microbiome Project (AMP), involving a population cohort of 127 elite soccer players in Serie A League, with well-defined phenotypes. This study was inspired by previous large-scale studies, which identified gut microbiome configurations and microbial taxa associated with inflammatory markers [24], blood lipids [25], and post prandial glucose responses [26]. In the AMP we characterized the composition and the function of the gut microbiome of 127 elite soccer players (age range 19-34 years). Figure 3 and 4.

Figure 3: Graphical summary of the Athlete's Microbiome Project (AMP)

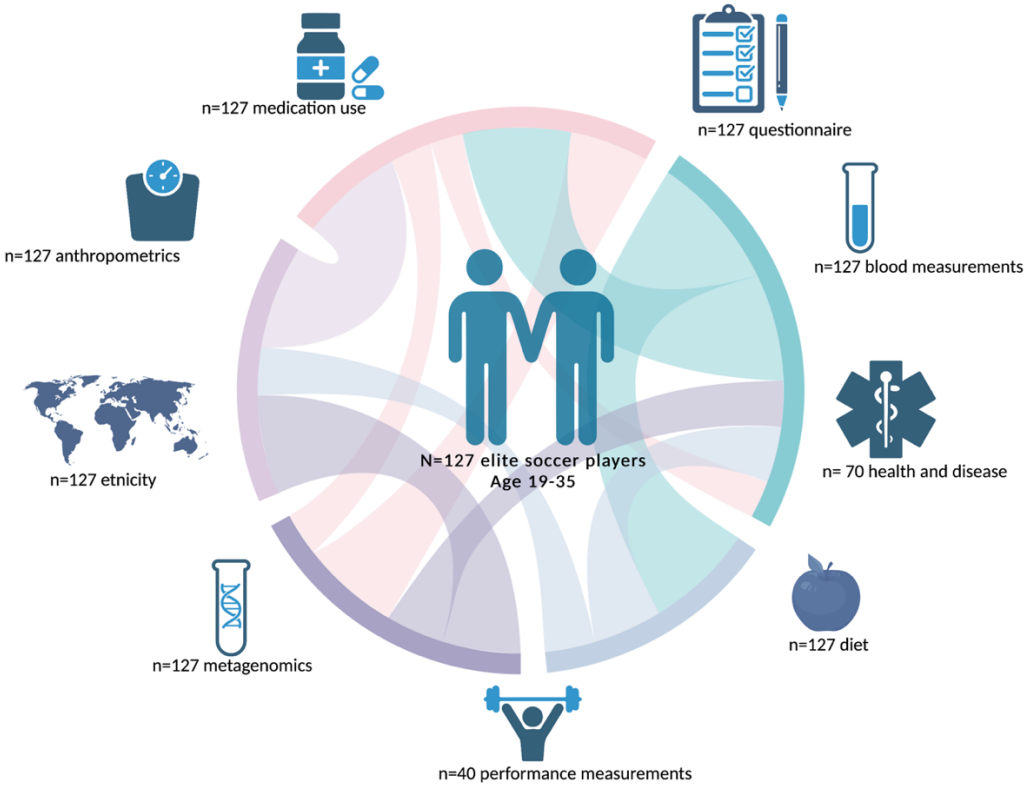


Figure 4. Project development

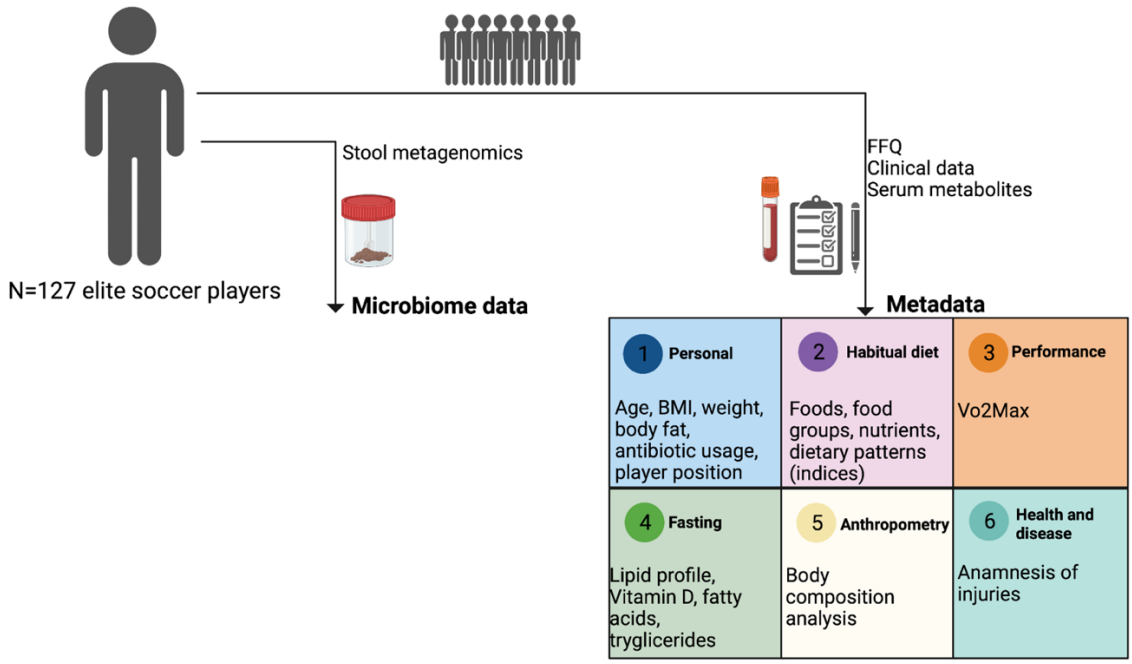


Figure 4: The AMP assessed the gut microbiome of 127 elite soccer players via metagenomic sequencing of stool sample. Phenotypic data obtained through in-person assessment and with the help of the medical staff, blood/stool collection and the return of validated study questionnaire queried a range of relevant host/environmental factors including personal characteristics (age, BMI, body composition, player position), habitual dietary intake using Food Frequency Questionnaire), fasting total lipid profile, anamnesis of injuries and performance test.

Main objectives:

-Identify microbial biomarkers related to the elite athlete population (less abundant and most abundant genera/species)

-Compare alpha-diversity indices before and after championship for paired-longitudinal samples

-Compare differences among different player position in term of taxonomy and alpha-diversity

-Compare the elite athlete's population against healthy control population

-Create an athlete's microbiome index (AMI)

-Investigate diet-gut microbiome interactions

Material and methods for data acquisition:

Population cohort and metadata collection

The Athlete Microbiome Project (AMP) cohort included 127 elite soccer players, members of the Italian Serie A league and recruited in 4 different teams (Genoa C.F.C., Fiorentina A.C.F, Hellas Verona, Parma Calcio). The sampling took place between September 2020 and May 2021, and all the participants signed a statement of informed consent before the beginning of the study.

A set of clinical data, including participant's birth date, height, weight, antibiotic use over the last month and habitual diet were collected at the enrollment. Fecal sample collection, blood analysis, anthropometry and performance measures were collected during the following 3 days at each training

facility. Briefly, athletes were asked to collect their fecal material (single defecation) in a plastic vial and place the vial in a labelled box. The nutritionist and/or the doctor of each team immediately collected the samples and delivered them within 72h to the research hub facility (University of Padova, Department of Biomedical Science).

Blood samples were drawn by nurses and analyzed by an independent certified clinical laboratory (https://www.synlab.it). Participants were asked to be fasted the morning of blood test.

Anthropometric measures, body composition analysis and performance measures Vo2Max (maximal oxygen consumption) were measured at each training center, in fasted state.

Microbiome sequencing and profiling

We performed deep shotgun metagenomic sequencing in stool samples from a total of 127 AMP participants. Computational analysis was performed using the bioBakery suite [27] of tool (http://huttenhower.sph.harvard.edu/biobakery) to obtain species-level microbial abundances for the taxa identified using the last version of MetaPhlan tool.

Microbiome sample collection

At each training facility, athletes were personally delivered a "microbiome kit pack" with the stool collection kit and information needed and ask to collect the stool sample within 3 days. We conducted a pre-study meeting with the medical staff of each team to explain the correct procedure and the 'best practices' of sampling, as accurately described by Segata et al.[28]. The sample collected, was deposited into a tube containing DNA/RNA Shield buffer (Zymo Research) to preserve the genetic integrity and expression profiles of samples at ambient temperatures (no refrigeration or freezing needed) and completely inactivates infectious agents (viruses, bacteria, fungi, & parasites). In this way, we facilitated the collection of the samples at each sport facility. Samples were stored at ambient temperature until returned to the research hub facility. Once receipted at university, and checked the sufficient biomass, the samples were sent by standard delivering to the laboratory

(https://igatechnology.com) to perform the wet analysis. The sample were homogenized, aliquoted and stored ay -80° in QIAGEN Power Beads 1.5mL tubes. The sample collection procedure was tested and validated internally comparing two different DNA extraction kits (Fast DNA and Zymo).

DNA extraction and sequencing

DNA was isolated by CELERO™ DNA-Seq kit (San Carlos, California, NuGEN Technologies, Inc.) using DNA/RNA Shield-fixed microbiome samples. Before library preparation and sequencing, the quality and quantity of the samples were assessed using Fragment Analyzer system (Agilent technologies). Only samples with a high-quality DNA profile were further processed. Both input and final libraries were quantified by Qubit 2.0 fluorometer (Termo Fisher) and quality tested by Agilent 2100 Bioanalyzer High Sensitive DNA assay. Libraries were then prepared for sequencing and sequenced on NovaSeq 6000 in paired ends *150 bp mode. At this step, the primary bioinformatic analysis included: 1.Base calling and demultiplexing. Processing raw data for both format conversion and de-multiplexing by Bcl2Fastq 2.0.2 version of the Illumina pipeline; and 2.Adapters masking. Adapter sequences are masked with Cutadapt v1.11 from raw fastq data using the following parameters: --anywhere (on both adapter sequences) --overlap 5 --times 2 --minimum-length 35 --mask-adapter. Lastly, a folder 'raw_reads' containing files with raw reads (R1: first read sequence; R2: second read sequence) and multiqc_report.html file, which aggregates results from primary bioinformatic analysis into a single report file with parameters that give insight into overall processing and sequencing quality, were provided.

Metagenome quality control and pre-processing

All sequenced metagenomes were quality control edited using the pre-processing process as specified in (https://github.com/SegataLab/preprocessing/). Pre-processing consists of three main step: (1)read-level quality control ; (2) screening of contaminants (host sequences) and (3) split and sorting of cleaned reads. The read-level quality control involves the removal of low-quality reads (i.e., quality

score QC < 20), fragment short reads (< 75bp) and reads with > 2 more ambiguous nucleotides. Contaminant DNA was identified using Bowtie 2 [29]. The sorting and splitting step created a standard forward (5' – 3'), reverse (3'-5') and unpaired reads output files for each metagenome.

Microbiome taxonomic profiling

The metagenomic analysis was performed following the general guidelines which relies on the bioBakery computational environment. The taxonomic profiling of organisms' relative abundances for each metagenomic sample was performed by MetaPhlan tool v. 4.0. (https://github.com/biobakery/metaphlan/wiki/). MetaPhlAn 4 relies on ~5.1M unique clade-specific marker genes identified from ~1M microbial genomes (~236,600 references and 771,500 metagenomic assembled genomes) spanning 26,970 species-level genome bins (SGBs: http://segatalab.cibio.unitn.it/data/Pasolli_et_al.html); 4,992 of them taxonomically unidentified at the species level. This set of species also included 83 species retrieved from defined by Co-Abundant Gene groups (CAGs) approach that were very genetically different from raw species represented by isolated genomes. (to make the idea clearer: .ost current approaches for analyzing metagenomic data rely on comparisons to reference genomes, but the microbial diversity of many environments extends far beyond what is covered by reference databases [30]). To estimate the microbiome species richness of everyone from the taxonomic profiles of AMP cohort, we computed alpha-diversity measures: the 'observed richness' (the number/count of different species found in the microbiome) and the 'Shannon index' (which combine the richness and the diversity). Microbiome dissimilarity (differences in the overall taxonomic composition between n samples) between athletes (beta diversity) was computed using Bray-Curtis dissimilarity on microbiome taxonomic profiles. The Bray-Curtis dissimilarity is always a number between 0 and 1. If 0, the two sites share all the same species; if 1, they don't share any species. It is based on count of each site: $BC_{ij} = 1 - \frac{2C_{ij}}{S_i + S_j}$ where

Where C_{ij} is the sum of the lesser values, for only those species in common between both sites. S_i and S_j are the total number of specimens counted at both sites.

Collection and processing of dietary information

Habitual dietary intakes were collected using Food Frequency Questionnaire (FFQ) -specifically the European Prospective Investigation into Cancer and Nutrition Norfolk Food Frequency Questionnaire (EPIC-Norfolk FFQ) (https://www.epic-norfolk.org.uk/for-researchers/ffq/). The questionnaire assessed the average intake in the past year.

Food Frequency questionnaire:

YOUR DIET LAST YEAR

For each food there is an amount shown, either a "medium serving" or a common household unit such as a slice or teaspoon. Please put a tick (✓) in the box to indicate how often, **on average**, you have eaten the specified amount of each food **during the past year**.

EXAMPLES:

For white bread the amount is one slice, so if you ate 4 or 5 slices a day, you should put a tick in the column headed "4-5 per day".

FOODS AND AMOUNTS	AVERAGE USE LAST YEAR								
BREAD AND SAVOURY BISCUITS (one slice or biscuit)	Never or less than once/month	1-3 per month	Once a week	2-4 per week	5-6 per week	Once a day	2-3 per day	4-5 per day	6+ per day
White bread and rolls								✓	

For chips, the amount is a "medium serving", so if you had a helping of chips twice a week you should put a tick in the column headed "2-4 per week".

FOODS AND AMOUNTS	AVERAGE USE LAST YEAR								
POTATOES, RICE AND PASTA (medium serving)	Never or less than once/month	1-3 per month	Once a week	2-4 per week	5-6 per week	Once a day	2-3 per day	4-5 per day	6+ per day
Chips				✓					

For very seasonal fruits such as strawberries and raspberries you should estimate your average use when the fruits are in season, so if you ate strawberries or raspberries about once a week when they were in season you should put a tick in the column headed "once a week"

FOODS AND AMOUNTS	AVERAGE USE LAST YEAR								
FRUIT (1 fruit or medium serving)	Never or less than once/month	1-3 per month	Once a week	2-4 per week	5-6 per week	Once a day	2-3 per day	4-5 per day	6+ per day
Strawberries, raspberries, kiwi fruit			✓						

The nutrient intakes were determined using FETA FFQ Software v.2.53 (https://www.epic-norfolk.org.uk/for-researchers/feta-download/) to calculate macro and micronutrients intake [31]. To better understand how changes in microbiome composition relates to diets, we decided to investigate

specific dietary indices based on *a priori* knowledge [32]. The following dietary indices were calculated: **(1)** Dietary intakes: Habitual dietary information was collected using FFQ and nutrient intakes were estimated by FETA software. Further, we used FETA to calculate the dietary indices; **(2)** HFD Index: the Healthy Food Diversity index consider the number, distribution and health value of the consumed food. This index is of importance because it considers simultaneously the number (count), the distribution (an individual with equal shares of foods has a greater food diversity than an individual who consumed 90% of 1 product and 10% of the others) and the health value (healthy or unhealthy products, such as fruits versus sugars) of a consumed food basket. The Health Value (HF) of foods were derived from the German Nutrition Society guidelines (https://www.dge.de/en./) and the weight of each food group was multiplied by its corresponding health value. Scores were divided by the maximum (health value 0,26) to bind values between 0 and 1 before multiplying with the Berry Index [33];

ORIGINAL_HFDI	UK FETA
Vegetables, fruits, leaf salads, juices	APPLES, AVOCADO, BANANAS, BEANS, BEANSPROUTS, BEETROOT, BROCCOLI, CABBAGE, CARROTS, CAULIFLOWER, COLESLAW, DRIED FRUIT, FRUIT JUICE, GARLIC, GRAPEFRUIT, GRAPES, GREEN BEANS, GREEN SALAD, LEEKS, LENTILS, MARROW, MELONS, MUSHROOMS, NUTS SALTED, NUTS UNSALTED, ONIONS, ORANGES, PARSNIPS, PEACHES, PEANUT BUTTER, PEARS, PEAS, PEPPERS, SEEDS, SMOOTHIES, SPINACH, SPROUTS, STRAWBERRIES, SWEETCORN, TINNED FRUIT, TOFU, TOMATOES, VEGETABLE SOUP, WATERCRESS,
Wholemeal products, Paddy	BROWN BREAD, BROWN RICE, CEREAL HIGH FIBRE, PORRIDGE, WHOLEM
Potatoes	BOILED POTATOES, CHIPS ROAST POTATOES, POTATO SALAD
White-meal products, peeled rice	MUESLI, NAAN POP TORTILLAS, WHITE BREAD, WHITE PASTA, WHITE RIC
Snacks and sweets - sugar, cakes, sweets, snack, pot	BISCUITS REDUCED FAT, CEREAL BARS, CEREAL SUGAR TOPPED, CHOCOLATE BARS, CHOCOLATE BISCUIT, CHOCOLATE DARK, CHOCOLATE MILK WHITE, CRACKERS, MARMITE, CRISPS, FIZZY DRINKS, FRUIT SQUASH, HOMEBAKED CAKE, HOMEBAKED SPONGE, JAM, KETCHUP, LOWCAL FIZZY DRINKS, PICKLES, PIZZA, PLAIN BISCUIT, QUICHE, READYMADE BUNS, READYMADE CAKE, READYMADE FRUIT PIES, READYMADE SPONGE, SAUCES, SAVOURY PIES, SUGAR, SWEETS, COCOA, COFFEE WHITENER, HORLICKS, HOMEBAKED BUNS
Fish, low-fat meat, low-fat meat products	CHICKEN, OILY FISH, ROE, SHELLFISH, WHITE FISH
Low-fat milk, low-fat dairy products	CHEESE REDUCED FAT, COTTAGE CHEESE, LOWFAT YOGURT
Milk, dairy products	CHEESE, DAIRY DESSERT, DOUBLE CREAM, FULLFAT YOGURT, ICE CREAM,
Meat products, sausages, eggs	BEEF, BURGER, CORNED BEEF, EGGS, FISH FINGERS, FRIED FISH, HAM, LA
Bacon	BACON
Oilseed rape, walnut oil	
Wheat germ oil, soybean oil	
Corn oil, sunflower oil	FRENCH, LOWCAL SALAD CREAM, OTHER DRESSING, SALAD CREAM
Margarines, butter	BUTTER, BUTTER REDUCED FAT, HARD MARGARINE, LOWFAT SPREAD, O
Lard, vegetable fat	

(3) (HEI) Healthy Eating Index: it assess to which extent an individual's food intake aligns with the dietary guidelines for American 2010, developed by the US department of Agriculture;

(4) PDIndex: plant based scoring system gives the consumption of plant foods positive marks and the consumption of any animal-derived foods negative markers. Unhealthy plant foods, such as refined grains and sugar and processed foods can also be given negative marks in what is termed an unhealthy plant-based diet index (uPDI) [34];

PDI Food Groups	UK_FETA		
Healthy			
Whole grain	BROWN BREAD, BROWN RICE, CEREAL HIGH FIBRE, PORRIDGE, WHOLEMEAL BREAD, WHOLEMEAL PASTA		
Fruits	BANANAS, DRIED FRUIT, GRAPEFRUIT, GRAPES, MELONS, ORANGES, PEACHES, PEARS, STRAWBERRIES, TINNED FRUIT, APPLES		
Vegetables	AVOCADO, BEANSPROUTS, BEETROOT, BROCCOLI, CABBAGE, CARROTS, CAULIFLOWER, COLESLAW, GARLIC, GREEN BEANS, GREEN SALAD, LEEKS, MARROW, MUSHROOMS, ONIONS, PARSNIPS, PEAS, PEPPERS, SPINACH, SPROUTS, SWEETCORN, TOMATOES, VEGETABLE SOUP, WATERCRESS, MARMITE, KETCHUP, PICKLES, SAUCES		
Nuts	NUTS SALTED, NUTS UNSALTED, PEANUT BUTTER, SEEDS		
Legumes	TOFU, LENTILS, BEANS		
Vegetable oils	FRENCH, OTHER DRESSING, SPREAD OLIVE OIL		
Tea and coffee	DECAFF COFFEE, FRUIT TEA, GREEN TEA, INSTANT COFFEE		
Less Healthy			
Fruit juices	FRUIT JUICE, SMOOTHIES		
Refined grains	MUESLI, NAAN POP TORTILLAS, WHITE BREAD, WHITE PASTA, WHITE RICE, CRISPBREAD, CORNFLAKES RICE KRISPIES, CRACKERS		
Potatoes	BOILED POTATOES, CHIPS ROAST POTATOES, POTATO SALAD, CRISPS		
Sugar sweetened beverages	FIZZY DRINKS, FRUIT SQUASH, LOWCAL FIZZY DRINKS, COCOA, COFFEE WHITENER, HORLICKS, HOT CHOCOLATE LOW FAT		
Sweets and desserts	BISCUITS REDUCED FAT, CEREAL BARS, CEREAL SUGAR TOPPED, CHOCOLATE BARS, CHOCOLATE BISCUIT, CHOCOLATE DARK, CHOCOLATE MILK WHITE, HOMEBAKED CAKE, HOMEBAKED		
Animal Food Groups			
Animal fat	BUTTER, BUTTER REDUCED FAT		
Dairy	CHEESE REDUCED FAT, COTTAGE CHEESE, LOWFAT YOGURT, CHEESE, DAIRY DESSERT, DOUBLE CREAM, FULLFAT YOGURT, ICE CREAM, SINGLE CREAM, MILK PUDDINGS		
Egg	EGGS		
Fish or seafood	OILY FISH, ROE, SHELLFISH, WHITE FISH, FISH FINGERS, FRIED FISH		
Meat	CHICKEN, BEEF, BURGER, CORNED BEEF, HAM, LAMB, LASAGNE, LIVER, MEAT SOUP, PORK, SAUSAGES, BACON		
Micellaneous animal based foo	LOWCAL SALAD CREAM, SALAD CREAM, PIZZA, QUICHE, SAVOURY PIES		
Co-variate			
Margarine	HARD MARGARINE, LOWFAT SPREAD, OTHER MARGARINE, POLYUNSATURATED MARGARINE, SPREAD CHOLESTEROL REDUCING, VERY LOWFAT SPREAD		
Alcohol	BEER, PORT, SPIRITS, WINE RED, WINE WHITE		
PDI Food Groups (18)	**PDI**	**hPDI**	**uPDI**
Healthy			
Whole_grain	+	+	-
Fruits	+	+	-
Vegetables	+	+	-
Nuts	+	+	-
Legumes	+	+	-
Vegetable_oils	+	+	-
Tea_and_coffee	+	+	-
Less Healthy			
Fruit_juices	+	-	+
Refined_grains	+	-	+
Potatoes	+	-	+
Sugar_sweetened_beverages	+	-	+
Sweets_and_desserts	+	-	+
Animal Food Groups			
Animal_fat	-	-	-
Dairy	-	-	-
Egg	-	-	-
Fish_or_seafood	-	-	-
Meat	-	-	-
Micellaneous_animal_based_fc	-	-	-

(5) The animal score: it categorized animal foods into "healthy" and "less-healthy"/"unhealthy" categories according to previous epidemiological studies (https://www.nhs.uk/live-well/eat-well/milk-and-dairy-nutrition/.); **(6)**aMED score: it measured the adherence to the Mediterranean diet by following the method outlined by Fung et al. [35].

Statistical analysis:

To test the association, Spearman correlations (reported with *p* in the text) were computed by the cor.test function from the stat R package version 3.5.1. Correlations and P values were computer for

the metadata (diet) and species, P values were corrected using false discovery rate FDR through the Benjamin–Hochberg procedure (BH step-down procedure), to avoid Type I errors (false positives), which are reported as q-value. We considered significant correlations with q<0.2, which means we should expect 2% of all the variables with q-value less than this to be false positives. To test the differential abundances among the groups (athletes versus control) we used 3 different algorithms: DESeq2, ANCOM and CODa4microbiome. The differential abundance analysis was carried out at genus level and the conditions were athletes and controls. Once identified several differences between the two groups, the criteria to consider the 'biomarkers' were: DESeq2: only Operational taxonomy Unit (OTUs) that reach genus level and FDR <0,001; COD4microbiome: only OTUs that reach genus level, ANCOM: only OTUs that reach genus level and detected -0.7=TRUE. Biomarkers whose log2FoldChange was between +0.5 and -0.5 were eliminated. Then we created the '*biomarkers.tsv*' file with the name of the OTUs that were found at least in 2 methods (*genera cross table*) and we added if these OTUs were abundant in athlete or controls. To calculate the Athlete Gut Index (AGI) for each sample, we created the function in R [36] and applied to our biomarkers (on both athletes and controls) :

$$DI = \left\{ \frac{\sum_{i=n} abundance(condition\ enriched)i}{n} - \frac{\sum_{j=m} abundance(control\ enriched)j}{m} \right\} \times 100$$

Then, we performed Wilcox.test to verify if indices were different among groups. Finally, we considered the distributions as normal distributions, and we calculated the z-scores for each sample. After that, Z-scores distribution were added to the plot, showing that there exists a different value of gut index for athlete when it used the biomarkers' group selected. This proved that the athlete's population was different from controls (different microbiome composition).

Results:

Population cohort

AMP project explores diet-exercise and microbiome interactions in a cohort of elite soccer players (n=127) with accompanying metadata covering anthropometrics, habitual dietary intakes, and demographic information (Methods). 127 shotgun metagenomic samples were collected, then sequenced and metagenomically profiled. Upon completion, participants will receive the following free of charge: results of their blood analyses and a "map" of their microbiota composition.

Microbial diversity and composition are associated with dietary patterns

We investigated whether specific microbial taxa might influence microbiome composition (microbial species responsible for diet-microbiome associations). We considered only the significant taxa (genera and/or species) that had abundance >2% in at least one sample and appeared in more than 10% of the samples. First, we noticed that the most significant associations included 2 dietary exposures, full-fat yogurt and milk with probiotic taxa such as *Bifidobacterium animalis* and *Streptococcus thermophilus* (p=0,045). However, given the potential nutrients-nutrients interactions, it was difficult to understand the independent association of other single food with specific microbial species. In this regard, and in accordance with Bowier et al. [32], and Johnson et al [37]., we identified understandable clusters of species associated with indices, such as the "healthy plant-based food" / "un-healthy plant-based food" (such as refined grains) as defined by the PDI. As a matter of fact, the genera linked to healthy plant-based food were predominantly butyrate producers, including *Roseburia hominis* and *Faecalibacterium prausnitzii*, whereas *Clostridia* taxa associated with less healthy plant-based food or animal-based food. These result support the concept that diet is a driver of microbiome variations, but the importance of food source (plant or animal-based food) and the food quality (unhealthy plant-based food) should be considered for improving microbiome and associated health outcome. (Figure 5 a,b)

Figure 5a.

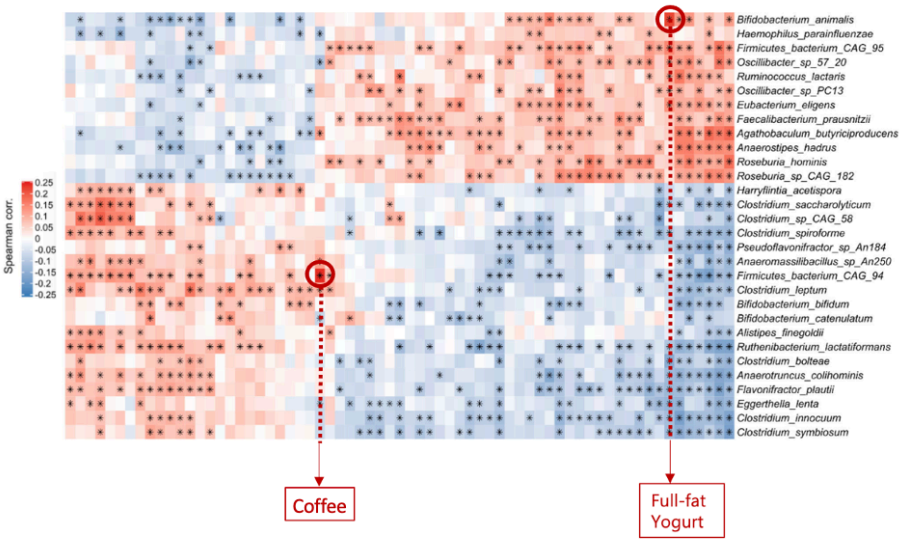
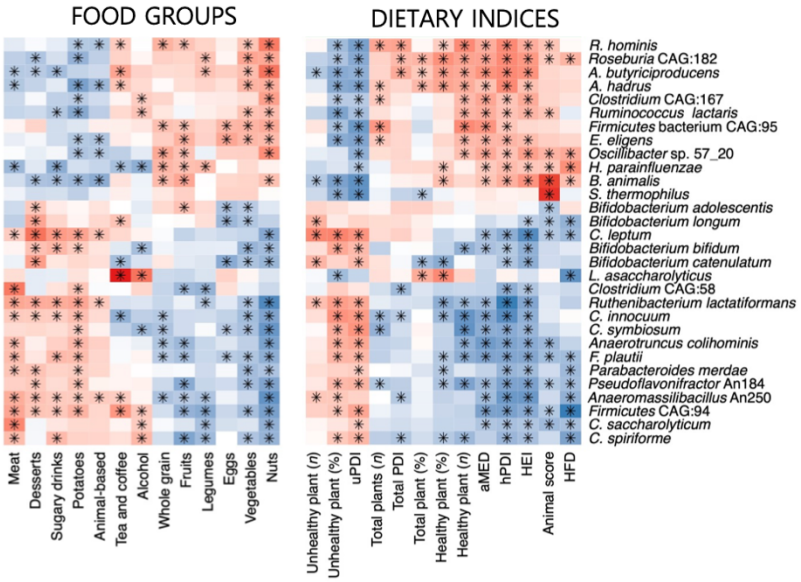


Figure 5b.



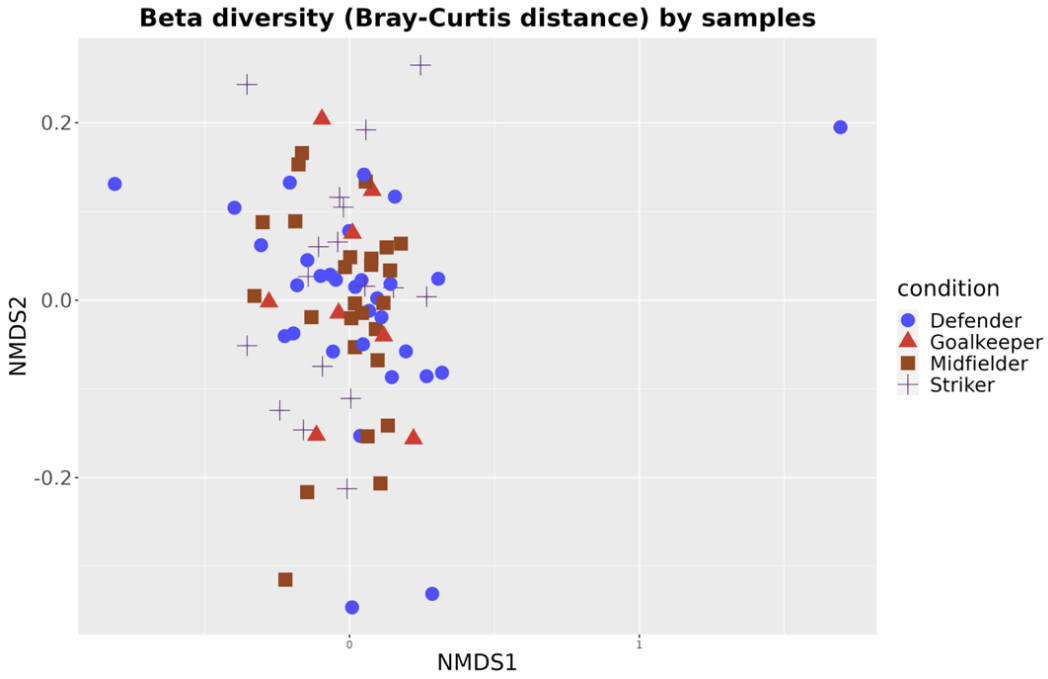
Comparison between elite soccer players and healthy control population

Our microbiota is composed of species that are found in high abundance and other in low abundance. This is because microbes undergo selective pressures from the host as well as from microbial competitors. After having an overview on the species diversity, to further understand whether elite soccer players might differently cluster, we plotted the beta diversity measure, that is, we clustered

the different species according to the taxonomic profile of their microbiota. For this purpose, we used the NMDS (non-metric multidimensional scaling) analysis with Bray-Curtis distances.

The first analysis demonstrated us that soccer players, divided by their different position into the field (striker, midfielder, defender and goalkeeper), did not cluster together. Figure 6.

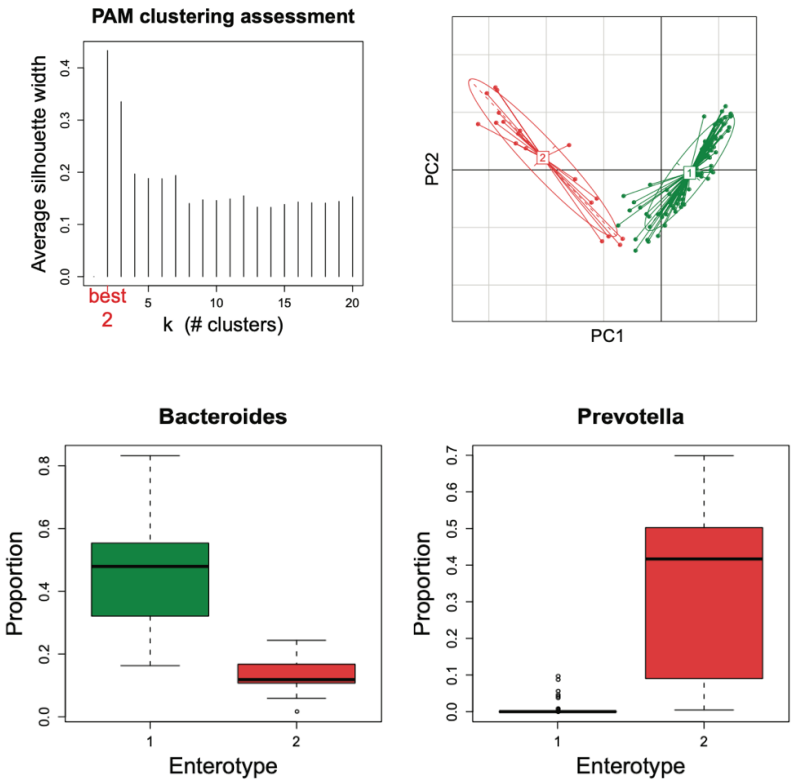
Figure 6.



Thus, following the suggestion by Aramugam et al.[38], we decided to investigate whether our AMP samples might partition into clusters based on Enterotypes* division, and we identified that *Prevotella* (enterotype 1) and *Bacteroides* (enterotype 2) mostly distinguished two clusters. Figure 7.

**Enterotype represent a classification of living organisms based on its bacteriological ecosystem in the gut microbiome*

Figure 7.



However, each of these enterotypes was affected by different microbial communities. To determine the phylogenetic basis of each enterotypes, we investigated in detail their differences in composition at species level. Enterotype 1 contained predominantly *Prevotella* and *Paraprevotella* (phylum Bacteroidetes) while Enterotype 2 was distinguished by the presence of *Alistipes* and *Parabacteroides* (phylum Bacteroidetes). However, the enterotype clustering was driven primarily by the ratio of the two dominant genera, *Prevotella* to *Bacteroides,* which defined the gradient across the two enterotypes.

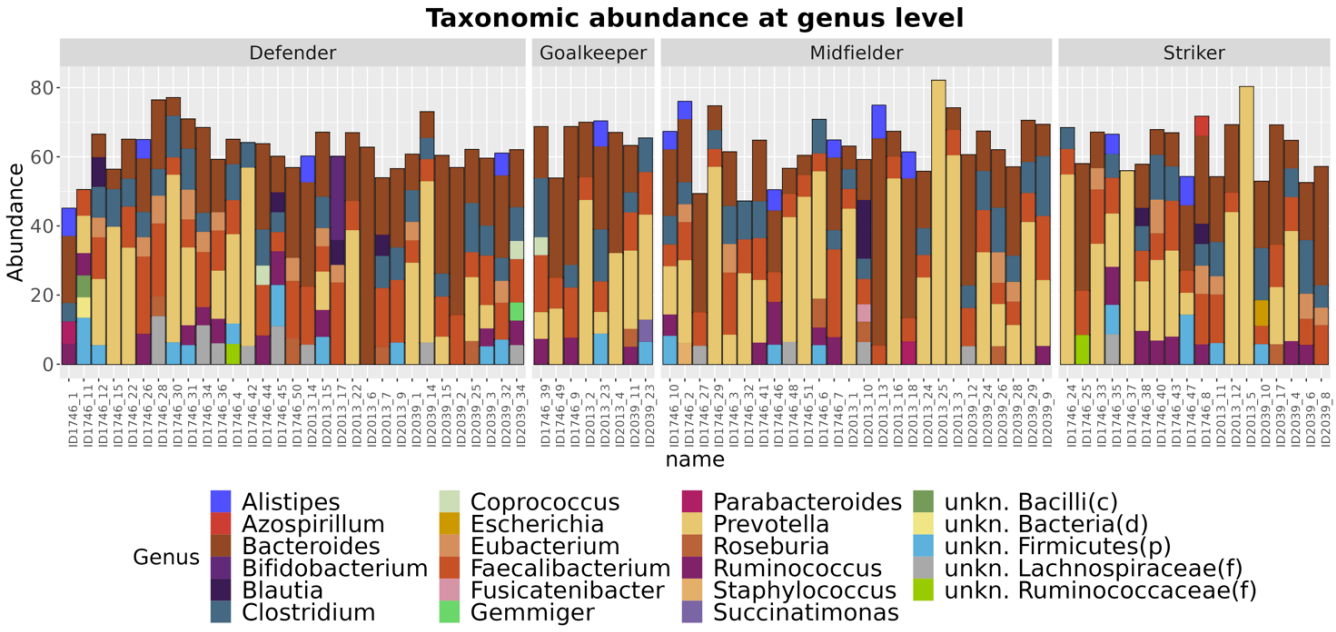
Finally, we decided to compare our athlete's cohort against a healthy population to investigate the differences in microbiome composition. The control population was composed by ten healthy Italian individuals, downloaded from NCBI SRA, ID project PRJNA331138. Considering that this is not the best way to proceed, because of the number of athletes is much higher than the controls, we discovered that, at genus level, there were some differences in the composition.

Specifically, we identified *Prevotella, Paraprevotella, Alistipes* more abundant in soccer players than in control group. Interestingly, among *Prevotella* genus, *Prevotella Copri* was the most abundant specie.

Microbial communities are not associated with different player position

To investigate whether the different player positions (striker, midfielder, defender, goalkeeper) might influence the gut microbiota composition, we calculated the relative counts of each genus and drew a taxonomy bar plot, divided by field position. The y-axis represented the percentage of genera within each sample (none reached the 100% because we eliminated all the genera that did not reach 5% to facilitate the interpretation and visualization of the data). Figure 8.

Figure 8.



Further, we performed the alpha diversity analysis by dividing the players by their positions. We presented alpha diversity in two ways. 1) Richness, that is the total number of different genres found in each sample (Figure 9) and 2) Shannon Index (Figure 10), that takes into account not only the number of different genera, but also the abundance within each sample of these.

Figure 9.

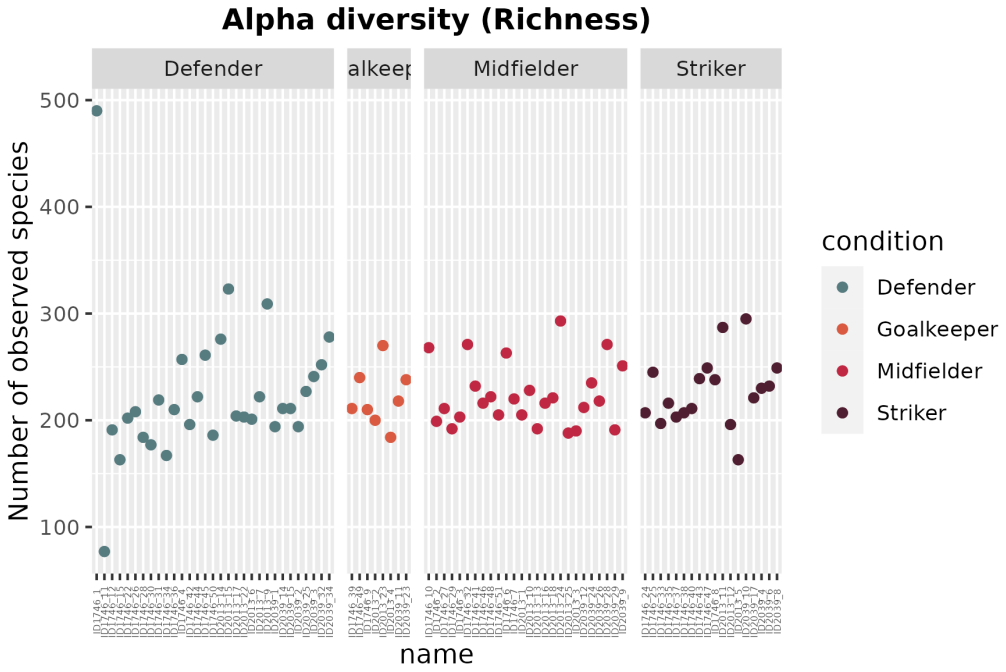
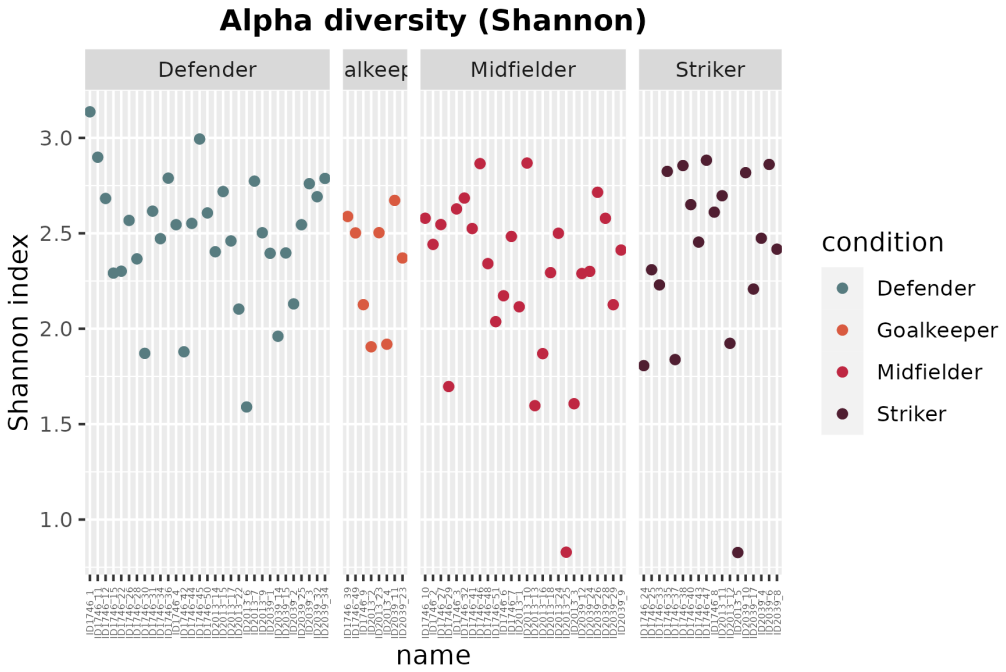
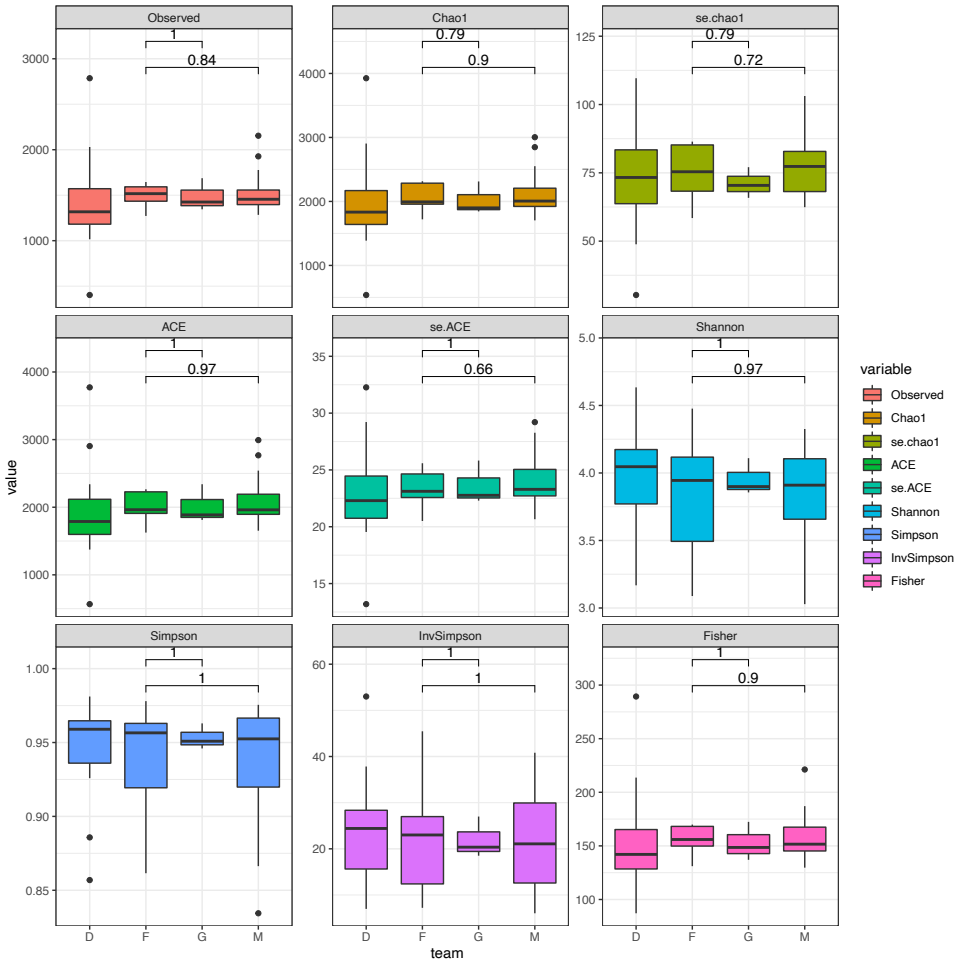


Figure 10.



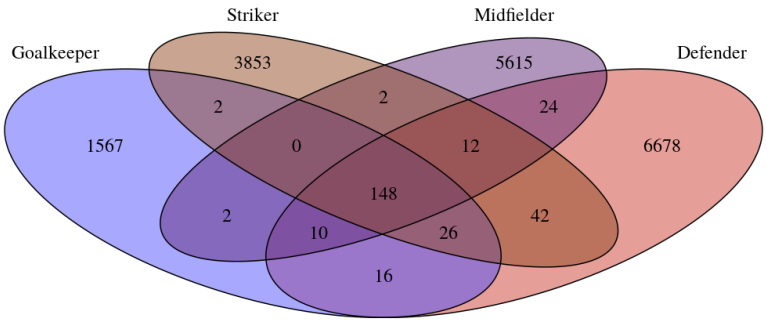
The results revealed that there was no statistically significant difference in alpha-diversity between different player position (D,F,G,M). Other indices of alpha diversity helped us to demonstrate the same result (Figure 11).

Figure 11.



We also presented the same results with a Venn diagram with the different genera found in each of the group. Figure 12.

Figure 12.

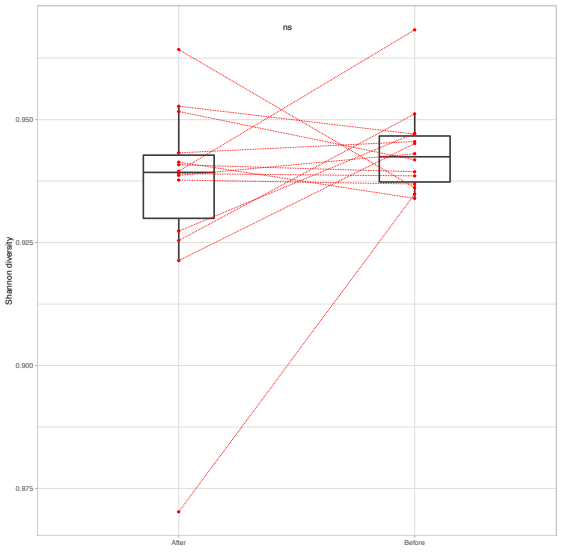


All together the results indicated that there were no differences in the gut microbiome of elite soccer players when split by player position.

Microbial community changes between pre - post competitive season

We decided to study the potential changes in gut microbiome during the competitive season. 10 paired samples were investigated. We noticed that the alpha diversity, measured as Shannon diversity, slightly decreased during the championship among these players (from September to March), (p=0,21). However, the analysis revealed the high inter and intra subject variability of the gut microbiome at baseline, and after 6 months of training. (Figure 13).

Figure 13.



These results demonstrate that under identical exercise training program, the inter-subjects and intra-individual microbiome variability is very high. Indeed, similarly to our findings, controlled feeding trials revealed that inter-subject microbiome variation is high even after period of identical dietary intake [39]. Our results showed that exercise-gut interactions are highly personalized and that the personalized response might be one reason for the small observed effect sizes of exercise and/or diet in shaping the gut microbiome in population level studies [40]. However, at compositional level, we found differences between pre and post championship for the genus *Prevotella* (p = 0.021, ES = 0.578), which decreased at the end of the season (Figure 14).

Figure 14.

Biomarker identification and Athlete's Microbiome Index (AMI)

A National Institutes of Health working group has defined a biomarker as "a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathological processes, or pharmacologic responses to a therapeutic intervention / treatment" [41].

At the beginning of the project, it was thought to identify microbial biomarkers related to the population 'elite soccer player' in order to potentially characterized this specific population and identified a more "health-associated" gut microbiota (In individuals without disease, "health-

associated" microbiota is preferred to the term "healthy microbiota", since gut microbial composition alone cannot predict any state of health or disease according to currently available research) (*Campbell K. Gut microbiota: interactive effects of nutrition and health, 2018*).

For this purpose, a statistical analysis on the abundance table was conducted (Methods) and a series of biomarkers were identified for the condition elite soccer players. At phylum level, we noticed that the elite soccer player cohort had a greater abundance of genera belonging to the genera *Prevotella*, whose presence was lower in healthy control population. Table 1, 2.

Table 1. Elite soccer player biomarkers (species level with adjusted p-value <0,01, log2Foldchange >0.5 or <-0.5 and average abundance among all samples >0.1%).

OTUs species BIOMARKER	**log2FoldChange**	**pValue**	**behaviour**
Porphyromonas uenonis	5,3307870	5,76E-04	most abundant
Paraprevotella clara	16,901574	5,39E-27	most abundant
Paraprevotella xylaniphila	15,154024	1,24E-22	most abundant
Prevotella amnii	11,415512	8,43E-138	most abundant
Prevotella bergensis	11,183919	4,04E-289	most abundant
Prevotella bivia	12,442637	8,17E-137	most abundant
Prevotella copri	21,165868	7,05E-131	most abundant
Prevotella corporis	12,077116	2,55E-131	most abundant
Prevotella dentalis	11,268097	1,09E-134	most abundant
Prevotella intermedia	14,487672	0,00E+00	most abundant
Prevotella ruminicola	12,467142	7,90E-112	most abundant
Alistipes putredinis	18,885677	0,00E+00	most abundant
Alistipes shahii	17,421591	0,00E+00	most abundant
Parabacteroides diastonis	16,770275	0,00E+00	most abundant

Parabacteroides golsteinii	13,845418	1,40E-167	most abundant
Parabacteroides Johnsonii	14,712837	3,64E-299	most abundant
Parabacteroides merdeae	17,610982	0,00E+00	most abundant

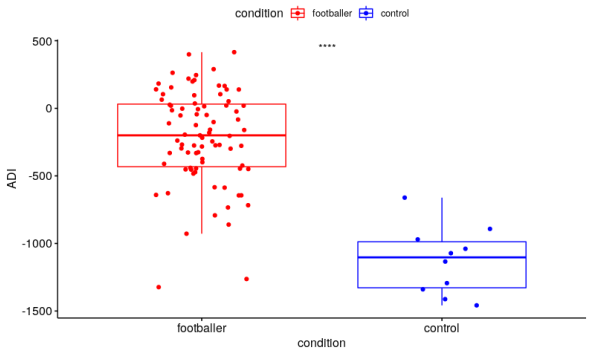
Table 2. Control biomarkers

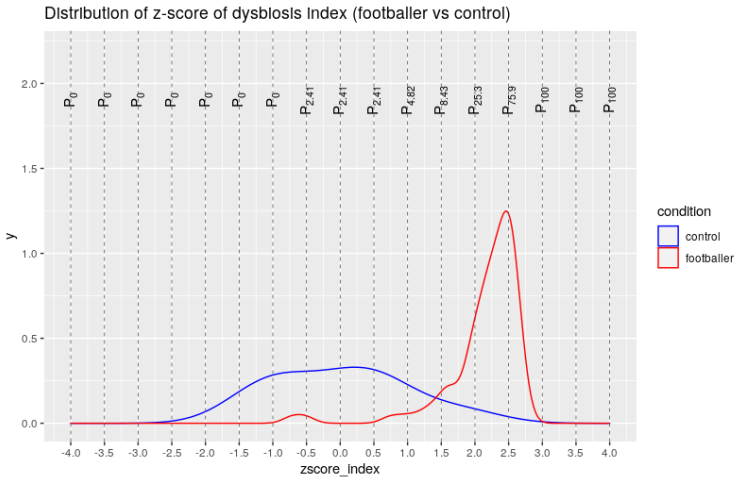
Bacteria	Bacteroidetes	Bacteroidia	Bacteroidales	Bacteroidaceae	Bacteroides most abundant
Bacteria	Firmicutes	Bacilli	Bacillales	Listeriaceae	Listeria most abundant
Eukaryota	Basidiomycota	Malasseziomycetes	Malasseziales	Malasseziaceae	Malassezia most abundant
Bacteria	Bacteroidetes	Bacteroidia	Bacteroidales	Odoribacteraceae	Odoribacte most abundant

Once the biomarkers were identified in elite soccer player cohort and the index calculated (Methods), box plot representing the distribution of the indices of the subjects, separated by the condition (elite soccer player and controls), were produced. The resulting plots showed that the indices were distributed in a different range among elite soccer players and controls (Figure 15), highlighting the potential of such index for athlete.

Figure 15. Box plot and distribution of indices

Figure 15.





Indeed, considering the specific index for 'soccer players', this proved that there could be the possibility of using such index to discern from "typical" to "atypical" associated-gut microbiome signature among soccer players. For example, the index may be used for the identification of an imbalance of gut microbiome in soccer players (our cohort of athletes was accurate selected to be "healthy", that means nor antibiotic or medication used at least 2 months before the sampling and, without gastrointestinal disorders that might have impact the results).

Discussion

The human gut microbiome is a fundamental indicator of the healthiness of a person, and it is shaped by many different factors such as diet and physical activity. Elite athletes often follow strict training and diet regimes to maximize performance and thus they can be used as a paradigm of the limit of the trained human body. After several years of intense training, elite athletes develop special features in terms of athletic performance but also in term of metabolic and physiological adaptations. For these reasons, AMP project was launched to investigate changes in elite athletes' gut microbiome composition and their relationship to habitual dietary intakes by collecting dietary data records and stool samples for shotgun metagenomics during the competitive season.

Our observational and longitudinal (pre-post championship) dataset allowed us to test whether diet-microbiome interactions were personalized. Notably, both microbiome signatures and dietary

components grouped into “health-associated” or “un-health-associated” clusters, in agreement with dietary quality and diversity indices (Methods). The identified food-microbiome interactions, such as those among favorable microbial clusters and the heterogeneity of food baskets (heathy diversity index) highlighted the importance of looking beyond “micro-macronutrients count” when investigating diet-microbiome interactions. In line with our findings, Johnson et al. showed that the conventional nutrients profiles alone to assess diet was insufficient to connect dietary intake with microbiome variations [37] because of the limited information available about nutrients and substrates (i.e., hundreds of additional chemical compounds present even in a single piece of fruit [42]) that are relevant to the microbiome. This fact might explain the inability to detect correlations between microbial functions and nutrient intake [40, 43] in some studies.

In addition, many other aspects might have obscured the ability to detect relationship in studies investigating the triad diet-exercise-microbiota. Indeed, the majority of the research presented pitfalls in:

-Food intake records: most of the questionnaire were not filled out by study participants and, in the FFQ, different type of plant-based food or protein sources were not usually documented;

-Microbial sequencing: 16S rRNA microbial data were often used to correlate to function (only ‘predicted’ functional profiling is possible). 16S rRNA is limited to identify bacteria at the genus level and not at species or strains levels;

-Bio-informatics analysis (not defined pipelines and standardized computational approaches);

-Small sample size;

-Involvement of “recreational athletes” incorrectly considered/defined as athlete (different physiological adaptation).

The consistence of our results resides in the quality of dietary recording (accurate FFQ provided by each nutritionist to each athlete) and analysis (food indices not only nutrients’ count), the large and well-defined sample size (127 elite soccer players with defined phenotypes) and the accurate metagenomic profiling pipeline (whole genome shotgun sequencing).

Another interesting aspect emerged from our research was that the microbiome was highly personalized among the same cohort of elite soccer players. This intra-subject variability might be one reason for the small effect size of diet and exercise in shaping the gut microbiome in some studies. In other words, the effect of exercise or diet might be much greater than these studies observed, but if the same food or training regime impact different bacterial populations in an individualized manner, such effects may not be totally detectable. However, our study deepened our understanding of the personalized diet-microbiome inter-relationships [26].

Finally, NMDS analysis and comparisons between elite athletes and control population showed that the microbiome of our elite athlete cohort was primary dominated by Prevotella and Bacteroidetes Enterotypes and, more specifically, at compositional level, by the species *Prevotella Copri*.

Prevotella is associated with plant-rich diets and its presence has been seen in healthy individuals, non-Westernized [44]. Indeed, non-Westernized population follow a traditional lifestyle and typically consume diets rich in fresh unprocessed food (vegetable and fruits). Although Westernization encompasses more factors and lifestyle modifications that diet alone, the association of *Prevotella* and (non)-Westernization may further support the hypothesis of diet being an important factor in selecting and shaping *Prevotella* populations. The higher prevalence of *Prevotella* in societies following a more traditional healthy-diet that the typical Westernized diet may also lend support for the health benefit of *Prevotella Copri*. Interestingly, and in line with our findings, Petersen et al. [45] identified associations between higher abundances of *Prevotella* and the amount of time that athletes (competitive cyclists) spent for exercising during an average week. Further, meta-transcriptomic analyses revealed that the increased abundance of *Prevotella* was correlated with a number amino acid and carbohydrate metabolism pathways, including branched-chain amino acid.

Al together these results provide a framework for a common constituent of the gut community in athletes who follow an exercise-rich lifestyle and habitual healthy dietary patterns (diet high in

complex carbohydrates, vegetable, fruits and high caloric intake). Overall, AMP project provides the first look into the gut microbiomes of elite soccer player. Indeed, this is the first large-study to identify a shared diet-microbial signature and segregating favorable and unfavorable taxa with specific dietary intakes. These data present opportunities for generating important hypotheses regarding how intense and long-lasting training influences the microbiome in elite soccer players. It is likely that multiple factors such as type of exercise, amount of exercise and diet influence how the microbiome of elite athletes are structured. Further studies will be important for understanding the impact of these factors on the metabolic capacity of the gut microbiome and how organisms such as *Prevotella Copri* may positively influence health and athletic performance. Finally, considering the characteristic of the microbiota to change from individual to individual, a personalized approach is required. As a resource, our findings will aid the utilization of gut microbiome in strategies for reshaping the microbiome to improve personalized dietary and training responses.

The resource: web platform developed for athlete's and health practitioner

From our results, and as recently demonstrated by Johnson et al., the gut microbiome is highly individual and microbiome responses to identical foods are strictly personalized.

These findings prompt us to develop a web platform, dedicated to athletes and health practitioners, able to deliver personalized nutritional advice, based on microbiome analysis. Indeed, since the microbiome composition and its metabolic contribution are related to health and performance, monitoring, and consequently modulating the microbiome, may represent a powerful strategy to improve athlete's health and performance. For such purpose, we created "*MICK 4Athletes*".

How does *MICK 4Athletes* works?

1) **Measure**: Collection of the stool sample and analysis of the gut microbiome ("gut microbiome profiling and analysis") + questionnaire for anamnesis + food frequency questionnaire.

We identified specific scores and insights for athletes, including inflammation status, microbial activity, richness and diversity, metabolic fitness, gut lining health, gut permeability.

2) **Nourish**: Based on microbiome data and metadata collected, a precise nutritional plan and/or supplement recommendations are delivered, such as:

minimizing certain foods ("you should still eat these foods but within specific limits"), avoiding specific foods and supplementation ("what you need and what you don't")

3) **Improve**: Following the recommendations could help to improve weight, inflammation, stress, sleep, energy levels, digestive wellness.

What do we decided to develop and include in the platform *MICK 4Athletes?*

1) **Identification of healthy bacteria**: certain bacteria are related to healthy metabolic profiles, *MICK 4Athletes* will identify the abundance of some microorganisms that are proven to help in different metabolic routes, such as carbohydrate metabolism 'for fueling' athletes
2) **Gut inflammatory status**: high intensity exercise together with some bacteria can increase the permeability of the gastrointestinal epithelial wall and diminish the mucus layer thickness which causes gut inflammation. *MICK 4Athletes* will help monitor the athlete's gut inflammatory activity.
3) **Gut Athlete's Index**: it will help to understand whether the athlete is in line with the " typical" pattern of microbial community of athlete or to identify the measure of the deviations from that score
4) **Vitamin synthesis**: some vitamins are produces in the gut by bacteria and have important roles in patient's health. *MICK 4Athletes* investigates the performance of the athlete's microbiome in producing vitamins and gives personalized food recommendation to improve it
5) **Microbiome diversity**: the higher diversity there is in the gut microbiome, the more resistant it is to invaders and the better it performs its metabolic roles. *MICK 4Athletes* provides an overall viewof the athlete's gut diversity and provides recommendations to increase it
6) **Production of short-chain fatty acids**: acetate, propionate, and butyrate are short-chain fatty acids that support the integrity of the gut lining and help prevent inflammation. They are produced from dietary diber in the gut microbiome. *MICK 4Athletes* provides an overview of the athlete's short chain fatty acid status and provides recommendations to improve its production
7) **Potential to break down nutrients**: *MICK 4Athletes* shows how well the gut microbiome breaks down nutrients such as fiber, protein, simple sugar and fat.

Each index is developed on specific algorithms. For example, the "**Gut inflammatory status"** considers the following:

- Mucin-degraders. If high, the score is 1. If on average, the score is 0. If low, the score is -1.
- Fiber intake. If high, the score is -1. If medium intake, the score is 0. If low, the score is 1.

- LPS-producers. If high, the score is 1. If on average, the score is 0. If low, the score is -1.
- SCFA-producers. If high, the score is -1. If on average, the score is 0. If low, the score is 1.

Each section has a score. In order to determine the inflammatory status, we summed the scores and we did the following:

(score mucin-degraders)*2 + (score fiber intake)*2 + (score LPS-producers) + (score SCFA-producers) = final score

If the final score is negative, the inflammatory status is LOW. If 0 or 1, INTERMEDIATE. If higher than 1, HIGH.

Note: Mucin-degraders and fiber intake are multiplied by 2 (more importance). There are 2 types of mucin-degraders: 1) mucin-degrading-specialists, and 2) mucin-degraders in general. The first group primary takes as carbon source the mucin available in the mucus layer (e.g. Akkermansia). The second group takes mucin as a carbon source when there is a shortage of other nutrients, such as fiber intake. For instance, Prevotella does not feed with mucin unless fiber intake is super-low. In this way, if mucin-degraders (in general) are high and fiber intake is low, the inflammatory status is high since it is expected that the mucus layer will be thinner.

The potential of *MICK 4Athletes?*

- **Actionable results**: it translates the microbiome analysis into multilevel comprehensive and customizable reports. Practitioners and researchers may interpret the results and customize the reports on-demand, in a few-click
- **Personalized health advice**: it integrates microbiome data with key athlete's metadata to provide actionable and personalized health insights
- **Automated and accurate**: it performs automated and real time analysis with the highest accuracy

- Trustable: it is developed on solid scientific basis and validated technology. The Mick (for general population) it is already chosen by hospitals and clinics in 5 countries.

Finally, we also improved our packaging to facilitate the collection and the process from sampling to analysis.

Next step: microbiome based solutions

Today, there is no doubt that the discovery of the gut microbiota community opened a promising and rapidly growing research field on the potential beneficial health effects of manipulating the gut microbiota. Indeed, the gut microbiota influences the function of the intestine and also metabolic tissue such as skeletal muscle. However, the majority of studies failed in the identification of "dysbiotic" or "healthy" microbiota profile, highlighting that microbiome is highly personalized and that food-based intervention aimed at modulating the intestinal community need to be tailored to the specific composition of an athlete's microbiome. For this reason, one major milestone would be to launch a larger-scale diet- microbiome interaction study applying a sportomics approach to "mimic the real challenges and condition that are faced during sport training and competition". Sportomics is defined as "non-hypothesis driven research on an individual's metabolite changes during sport and exercise". Such an integrated approach will open the door to personalized nutrition/training program based on microbial composition. It could lead to microbiome-based solutions for health or performance by helping the design of new supplements and probiotics. In addition to the new monitoring application, this strategy could lead to optimized diets through personalized nutrition based on an individual's microbiome make up and workout intensity. Finally, based on the previous large-scale studies, it would be interesting to measure and understand the unique metabolic responses to food or supplements in elite athletes. For example, it may be of interest to investigate how blood sugar and fat levels change after we eat or, how metabolites that are biomarkers of microbial metabolism of specific amino acids can change after the intake of different protein sources or how

fiber and prebiotics can be better tolerated by athletes. These studies should also include some relevant measurements of performance and recovery (i.e., training periodization, recovery techniques, sleep) to develop a global vision of body adaptation through exercise.

PART 2 - STUDY 2

Acute cocoa supplementation influences long-chain n-3 polyunsaturated fatty acids levels and microbiota composition in elite soccer athletes: introduction

See our systematic reviews for the general background:

Sorrenti V, Ali S, **Mancin L**, Davinelli S, Paoli A, Scapagnini G. Cocoa Polyphenols and Gut Microbiota Interplay: Bioavailability, Prebiotic Effect, and Impact on Human Health. Nutrients. 2020 Jun 27;12(7):1908. doi: 10.3390/nu12071908. PMID: 32605083; PMCID: PMC7400387

Increasing evidence demonstrate that consumption of dark chocolate is associated to lower risk of cardiovascular mortality due to the high content of cocoa' flavanols [46-48].

Both *in vitro* and *in vivo* studies showed that cocoa flavanols promote anti-inflammatory effects.

Mao et al., demonstrated that human immune cells incubated with cocoa flavanols exerted ani-inflammatory effect through inhibiting mRNA expression of pro-inflammatory cytokines, including interleukin IL-6 and tumor necrosis factor TNFα [49, 50]. Accordingly, in humans, acute flavanol-rich dark chocolate exerted anti-inflammatory effects both by increasing mRNA expression of the anti-inflammatory cytokine IL-10 and by attenuating stress reactivity of the pro-inflammatory transcription factor NF-kB and of the pro-inflammatory interleukin IL-6 at mRNA-levels [51].

More recently, it has been demonstrated that short-term intake of a flavanol-rich cocoa improved lipid profile and positively modulate the metabolism of polyunsaturated fatty acids (PUFA) in a cohort of healthy individuals [52].

The omega-3 and omega-6 polyunsaturated fatty acids (PUFAs) are among the most studied nutrients in human metabolism since the effects of these essential fatty acids are clinically relevant [53].

Indeed, the anti-inflammatory effects of PUFAs have already been harnessed in various chronic inflammatory condition and autoimmune diseases. Moreover, beyond the well-documented health benefits of PUFAs in cardiovascular health, weight management, and blood sugar control, these fatty acids play an important role in muscle strength, in reducing muscle inflammatory processes caused by intense exercise, and injury prevention in athletes.

However, these fatty acids do not have the same effect. It is known that omega 6 and omega 3 have opposing properties.

An altered fatty acid profile characterized by an excess of omega 6, creates a pro-inflammatory environment associated with inflammatory processes [54, 55], while, long-chain n-3 PUFAs decrease the production of inflammatory mediators (eicosanoids, cytokines, and reactive oxygen species), give rise to anti-inflammatory mediators (resolvins) [54]. As such, they may be of therapeutic use in a variety of acute and chronic inflammatory settings or adopted for the enhancement of muscle recovery[56, 57].

Of note, the anti-inflammatory efficacy of omega-3 may be improved if intakes of omega-6, especially arachidonic acid, are decreased [54].

Although there are several different omega-3 fatty acids, eicosapentaenoic acid (EPA), docosahexaenoic acid (DHA) and alpha-linolenic acid (ALA) are most prominent in term of human physiology and metabolism. However, while it is possible for ALA (found in plant-based source) to be converted to EPA and DHA in the body, the rate of conversion is modest [58]. Thus, consumption of EPA and DHA (found primarily in marine-based sources) is the best strategy to attain these nutrients. Omega-3 fatty acid status can be determined via dietary assessment (targeted food frequency questionnaire) o via blood markers. It is noteworthy that the arachidonic acid (AA): EPA ratio has been suggested as a potentially more relevant indicator of balance between omega-6 and omega-3 fatty acids in the diet since arachidonic acid and EPA compete metabolically for eicosanoid

production [52, 59]. Of note, most athletes have been shown to consume sub-optimal omega-3 fatty acid, [60] thus highlighting the need to identify additional nutritional strategies to reach the optimal omega-3 fatty acid status throughout the season. Moreover, athletes may likely require more omega-3 than the general population with factors such as energy metabolism, training volume and the inflammatory response to exercise all influencing needs [59, 61].

Several dietary recommendations exist. Algae, phytoplankton, and other marine microorganisms are natural producers of EPA and DHA, and thus, in turn, fish and seafoods that consume these microorganisms are the richest sources in the food supply.

Interestingly, also plant-derived polyphenols, such as cocoa flavanols, report a "fish-like effect" adequate for increasing blood EPA and DHA levels, and improving (AA): EPA ratio [52, 62-64].

Indeed, cocoa and chocolate products have much higher flavonoid concentration per weight than other flavonoid food sources, such as red wine, green tea or black tea [65]. Among milk and dark chocolate, the latter contains higher amount of cacao and higher amount of flavonoids (i.e., 951 mg polyphenols /40g serving compared to 394 mg milk chocolate) [65].

One mechanism by which cocoa polyphenols may influence the balance between omega-6 and omega-3 relates to the gut microbiota activity[5]. Indeed, once cocoa polyphenols reach the intestine, can interact bidirectionally with the gut microbial community. The cocoa polyphenols modulate the intestinal microbiota by exerting prebiotic effects, while the microorganisms metabolize the polyphenols in bioactive compounds. The bioactive molecules can enter the circulation, reach the target organs, and exhibit their activities [66]. In addition, a functional physiological gut microbiota (i.e., stability over time) is essential for athletes' health and performance since it represents a key player along the signaling brain-gut axis [67] and skeletal muscle-gut axis [3].

To the best of our knowledge, no previous interventional studies have examinate the effect of high flavanol dark chocolate on the metabolism of PUFAs in elite athletes, and particularly its effect on AA:EPA ratio and gut microbiota composition. Interestingly, AA:EPA ratio has been recently identified as a biomarker of running-related injuries [59].

For this reason, we conducted a dietary intervention study to assess the impact of 30g of dark chocolate on clinical biomarkers of omega-3 fatty acid status and to test whether dark chocolate supplementation may affect the athlete's gut microbial community. Previous work investigating gut microbiota in athlete's population has numerous confounding factor such as age, lifestyle, diet, and different physical activity level. This study focuses on a homogeneous group of elite-soccer players who shared baseline dietary habits and physical activity level. Recruiting such a well-controlled cohort enable us to highlight changes in plasmatic biomarkers and microbial features associated specifically with dark chocolate intervention. Critically to the aim of this study, diet and supplementation were rigorously controlled which enabled data to be attributable to the treatment.

Material and methods

Participants

We performed a randomized controlled trial in a cohort of 38 elite football players (27.22 ± 4.31 years) during the last month of the regular season (May 2022). Exclusion criteria were any disease or serious medical condition, probiotic use, special diet such as vegan or vegetarian, and the use of non-steroidal anti-inflammatory drugs during the previous 7 days. Players were member of the Italian first league (Serie A), Genoa C.F.C. and Fiorentina A.C.F. The subjects were engaged in a ~120-minute training, 6 times/week and 90 minutes, match/week. The training includes a 15-minute warm-up, 30 minute technical/tactical skills, 30 minutes of strength training and 15-minutes cool down. The training program schedule and the number of matched played were the same among the two teams. The participants were instructed to maintain their habitual diet and supplementation, according to their anthropometric characteristics and to the amount of calories coming from chocolate intake. All the athletes, medical staff and the head of coach were explained the study's purposes, risk, and benefit, and they gave a written informed consent. Further, the nutritionists were asked to check with

a daily frequency the athletes' diet in order to control the maintenance of their habitual dietary intakes during the study (both quality and quantity of food.

Study design

Elite football players were randomly divided into two groups: treatment group provided with 30g dark chocolate 88% in tablet (DC n=19) and control group with 30g white chocolate (WC n=19) (0 mg flavanols). The participants were instructed to refrain from consuming chocolate or flavanol-containing products (i.e., blueberry, strawberry, blackberry, green tea, red wine, and pomegranate juice), starting from 3 days before the beginning of the study and throughout the intervention period. The dark chocolate and white chocolate were provided every morning to each subject by the nutritionist of each team, while, during off days, the chocolate was portioned into individual serving sizes and provided to each player. For each assessment, participants arrived at the training center in the morning, after an overnight fast. Before the beginning and the end of the study, a 3-days food record (3DR) was recorded (2 weekdays and 1 week-end day) by the nutritionist, blood and fecal sample was collected (between 08:00 and 09:00 hours), and anthropometry measures were determined. Briefly, for stool samples, athletes were asked to collect their fecal material (single defecation) in a plastic vial and place the vial in a labelled box. The nutritionist and/or the doctor of each team immediately collected the samples and delivered them within 72h to the research hub facility (University of Padova, Department of Biomedical Science).

Randomization and sample size

Randomization was conducted using an online computer-generated sequence (https://www.graphpad.com), matched for BMI. Sample size calculation was computed with G*Power [68] from similar studies with healthy individuals who consumed flavanol-rich cocoa.

Plasma Epicatechin extraction and quantification by HPLC-FLD/UV

Plasma samples were extracted by the method described by Spadafranca et al. [69]. Epicatechin was hydrolyzed enzymatically using beta-glucuronidase and sulfatase and subsequently extracted by addition of 1ml acetonitrile, and the mixture was centrifuged at10 000 g for 5 min at 4°C. After centrifugation, 50 µl supernatant was injected into the HPLC column for separation, detection, and analysis. The HPLC analysis was performed using an HPLC system (Agilent 1200 Infinity Series HPLC system, Santa Clara, USA). Spectroscopy data from all peaks were captured in the range of 210–400 nm, and chromatograms were recorded at 279 nm. Separations were carried out at a flow rate of 1·5 ml/min. The identification of the epicatechin was made by comparison of retention times and spectra with those of commercially available standard compound (−)-epicatechin, E-1753, Sigma).

Blood biochemical analysis

Blood samples were collected using EDTA tubes and left for 10–15 min at room temperature. To obtain plasma aliquots samples were centrifuged (1800 × g, 15 min, 4 °C) and stored at - 80 °C until analysis. Total cholesterol, HDL cholesterol, low-density lipoprotein (LDL) cholesterol, and triglycerides were measured using colorimetric enzymatic tests. The plasma concentration of oxidized LDL (oxLDL) was quantified with a specific ELISA kit (Immundiagnostik AG, Bensheim, Germany).

Fatty acid analyses

Blood was collected into heparinized evacuated tubes and centrifuged at 1.000 × g for 10 min. Total lipids were extracted from plasma with chloroform: methanol:0.90% KCl (2:1:0.2, v/v/v). Plasma fatty acid composition was determined from 2 mL of the lipid extract after transformation into isopropyl esters. Separation of isopropyl esters was done on a gas chromatograph using a capillary column (internal diameter: 0.32 mm). Column conditions were 175 °C for 4 min and then increased by 3 °C/min to 220 °C for 30 min. Helium was used as the carrier gas (flow rate: 2 mL/min). The

peaks were identified by comparison with reference fatty acid esters and peak areas were measured with an automatic integrator. The results for each fatty acid were expressed as a percentage of total fatty acids. The AA-to-EPA, AA-to-DHA ratios were calculated. We analyzed fatty acids that were greater than 0.01% of peaks detected.

Microbiome sequencing and profiling

We performed deep shotgun metagenomic sequencing in stool samples from a total of 38 elite soccer players. Computational analysis was performed using the bioBakery suite of tool (http://huttenhower.sph.harvard.edu/biobakery) to obtain species-level microbial abundances for the taxa identified using the last version of MetaPhlan tool.

Microbiome sample collection

At each training facility, athletes were personally delivered a "microbiome kit pack" with the stool collection kit and information needed and ask to collect the stool sample within 3 days. The sample collected, was deposited into a tube containing DNA/RNA Shield buffer (Zymo Research) to preserve the genetic integrity and expression profiles of samples at ambient temperatures (no refrigeration or freezing needed) and completely inactivates infectious agents (viruses, bacteria, fungi, & parasites). In this way, we facilitated the collection of the samples at each sport facility. Samples were stored at ambient temperature until delivered to the laboratory (https://igatechnology.com) to perform the wet analysis. The sample were homogenized, aliquoted and stored ay -80° in QIAGEN Power Beads 1.5mL tubes. The sample collection procedure was tested and validated internally comparing two different DNA extraction kits (Fast DNA and Zymo).

DNA extraction and sequencing

DNA was isolated by CELERO™ DNA-Seq kit (San Carlos, California, NuGEN Technologies, Inc.) using DNA/RNA Shield-fixed microbiome samples. Before library preparation and sequencing, the

quality and quantity of the samples were assessed using Fragment Analyzer system (Agilent technologies). Only samples with a high-quality DNA profile were further processed. Both input and final libraries were quantified by Qubit 2.0 fluorometer (Termo Fisher) and quality tested by Agilent 2100 Bioanalyzer High Sensitive DNA assay. Libraries were then prepared for sequencing and sequenced on NovaSeq 6000 in paired ends *150 bp mode. At this step, the primary bioinformatic analysis included: 1.Base calling and demultiplexing. Processing raw data for both format conversion and de-multiplexing by Bcl2Fastq 2.0.2 version of the Illumina pipeline; and 2.Adapters masking. Adapter sequences are masked with Cutadapt v1.11 from raw fastq data using the following parameters: --anywhere (on both adapter sequences) --overlap 5 --times 2 --minimum-length 35 --mask-adapter. Lastly, a folder 'raw_reads' containing files with raw reads (R1: first read sequence; R2: second read sequence) and multiqc_report.html file, which aggregates results from primary bioinformatic analysis into a single report file with parameters that give insight into overall processing and sequencing quality, were provided.

Metagenome quality control and pre-processing

All sequenced metagenomes were quality control edited using the pre-processing process as specified in (https://github.com/SegataLab/preprocessing/). Pre-processing consists of three main step: (1)read-level quality control ; (2) screening of contaminants (host sequences) and (3) split and sorting of cleaned reads. The read-level quality control involves the removal of low-quality reads (i.e., quality score QC < 20), fragment short reads (< 75bp) and reads with > 2 more ambiguous nucleotides. Contaminant DNA was identified using Bowtie 2. The sorting and splitting step created a standard forward (5' – 3'), reverse (3'-5') and unpaired reads output files for each metagenome.

Statistical analysis

Lipid profile and AA:EPA ratio

To each outcome, change from baseline was calculated for each time point with an independent t test. Results are presented as the mean ± SDs. We used an estimated standard deviation of 0.5 and the two tailed alpha set at 0.05. The two-way ANOVA was performed to identify differences between baseline and endpoint (at week 4) and considering treatment (dark chocolate, white chocolate) as inter-subjects factors. Statistical analyses were performed using SPSS Statistic software version 22 (IBM Corp., Armonk, NY, USA). Post hoc analyses were performed using Bonferroni test. In addition, effect size (ES) calculation was done with Cohen's *d*. The normal Gaussian distribution of the data was verified by the Shapiro-Wilk test.

Microbiome taxonomic profiling

The metagenomic analysis was performed following the general guidelines which relies on the bioBakery computational environment. The taxonomic profiling of organisms' relative abundances for each metagenomic sample was performed by MetaPhlan tool v. 4.0. (https://github.com/biobakery/metaphlan/wiki/). Microbiome dissimilarity (differences in the overall taxonomic composition between samples) between athletes (beta diversity index) was computed using Aitchison distance between two observations (microbiome profile pre and post intervention). Once measures of beta diversity have been calculated, the entire data set has been visualized using the ordination method of PCoA. PCoA method is used to explore and visualize dissimilarities of data by performing a rotation of the inter-sample distance matrix in order to represent those distances as accurately as possible in a small number of dimensions. Permutational Analysis of Variance (PERMANOVA) was used to test which variables = metadata_soccer[,c("ethnicity","team_location","timepoint","treatment","player_name")] were most significant in explaining the % of variance. To investigate the differences pre-post between the two groups (DC vs. WC) in the relative abundances at genus and species levels, we used the adonis R function, and post-hoc comparisons with a paired Wilcoxon test with FDR correction.

Microbiota community variation explained by metadata variables

The contribution of metadata variables (player_name, treatment, timepoint, team location, ethnicity) to interindividual microbiota community variation was determined by PERMANOVA analysis on genus-level Aitchison dissimilarity with the functions in the vegan R package. The cumulative contribution of metadata variables was determined by forward model selection on PERMANOVA with the ordiR2step function in vegan, with variables that showed a significant contribution to microbiota community variation ($P_{adj} < 0.05$) in the previous step.

Microbiome and metadata associations

Taxa unclassified at the genus level or present in less than 10% of samples were excluded from the statistical analyses. Wilcoxon paired tests were used to test the differences of continuous variables between two different groups

Results

Characteristic of study population

This study was conducted in May 2022. Of the forty-two athletes who were enrolled into the study, thirty-eight were eligible and randomly assigned to one of the two groups. All thirty-eight athletes completed the intervention period, and all data were included in the analysis. No significant differences were noted over the intervention period for anthropometric parameters (weight, BMI and percent of body fat) and macronutrients intakes across the groups. Differently, plasma total polyphenols resulted to be unchanged in WC group (from 150.4± 23.2 μg GAE/ml to 147.1±34.4 μg GAE/ml), while the content increased in DC group (from 154.7± 18.6 μg GAE/ml to 185.11±57.6 μg GAE/ml).

Dietary nutrition intake

There were no differences in dietary nutrient intakes between groups at baseline and after 4 weeks of intervention according to analysis of diets records.

Effects of dark chocolate on plasma lipid profile

Compared with baseline, 4 weeks of dark chocolate intervention revealed a statistically significant decrease in total cholesterol in DC group (-32.47± 17.18; -18.15%; P<0,001; time* treatment interaction P=0,032). No significant changes were observed for total plasma cholesterol in the WC group. Moreover, the DC group showed a reduction in triglycerides (-6.32± 4.96; - 7.62%; P<0,001; time* treatment interaction P=0.392) and LDL cholesterol (-18.42 ± 17.13; -13.67%; P<0,001; time* treatment interaction P=0.195) after intervention, compared with baseline. Differently, control group did not show any significant change in the level of triglycerides or LDL cholesterol. Furthermore, we demonstrated that 4 weeks of dark chocolate increased the level of HDL (3.26 ± 4.49; 6.63%; P=0,005; time* treatment interaction P=0.44). No differences were observed in DC group.

Effects of dark chocolate on AA/EPA ratio

Four weeks of dark chocolate supplementation significantly decreased the AA/EPA ratio. Indeed, daily supplementation of 30g of dark chocolate was associated with a significant reduction in AA/EPA ratio (-5.26 ± 2,35; -54.17%; P=<0,001; time* treatment interaction P=0.078). A weak reduction in the AA/EPA ratio was observed in the control group (-0.47 ± 0.73; -6.41%; P=0.012), suggesting the efficacy of dark chocolate on AA/EPA ratio.

Greater microbial taxonomic stability is associated with dark chocolate intake

Athletes were randomly divided into dark chocolate group intervention and white chocolate group control. Subject variance, based on Aitchison dissimilarity distance, was used as a measure of microbial stability within individual participants (Figure 16). Microbial communities of athletes

enrolled in dark chocolate group were more stable over time exhibiting lower within-subject community dissimilarity than subjects of control group (wilcoxon non paired, two.sided, p-value = 0.0001226) (Figure 17). Conversely, athletes in the control group showed a greater variation in gut microbial communities across the study. There results suggest that dark chocolate intake might have promoted a greater gut microbial stasis in athletes.

Figure 16. PCoA based on Aitchison dissimilarity distance

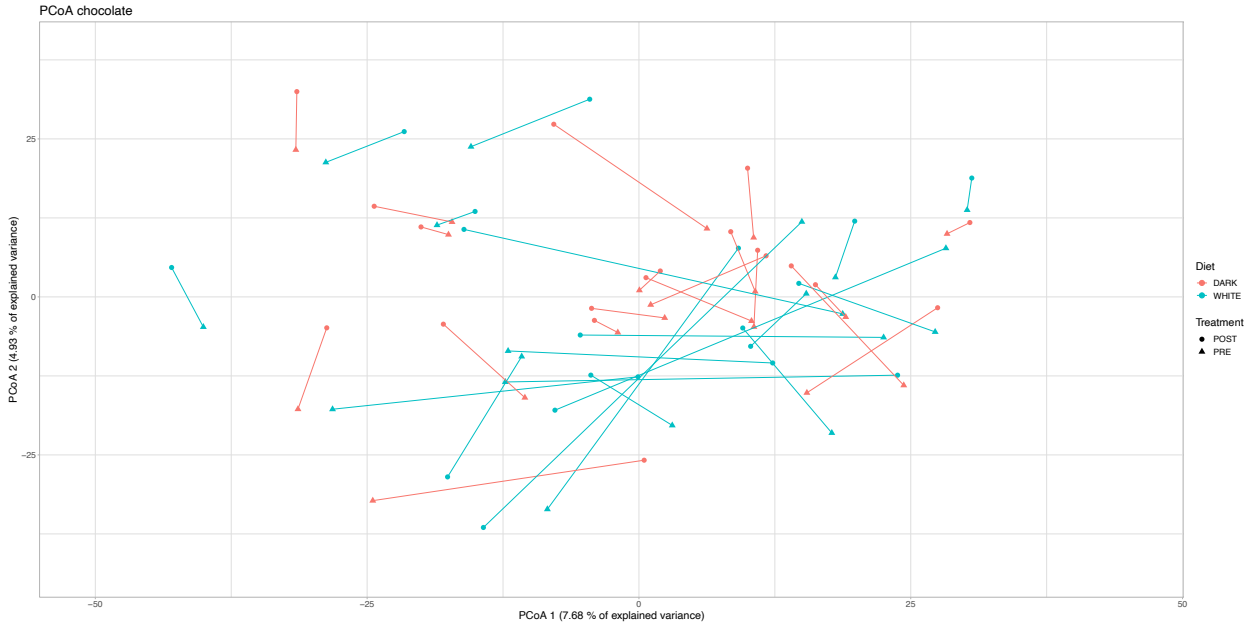
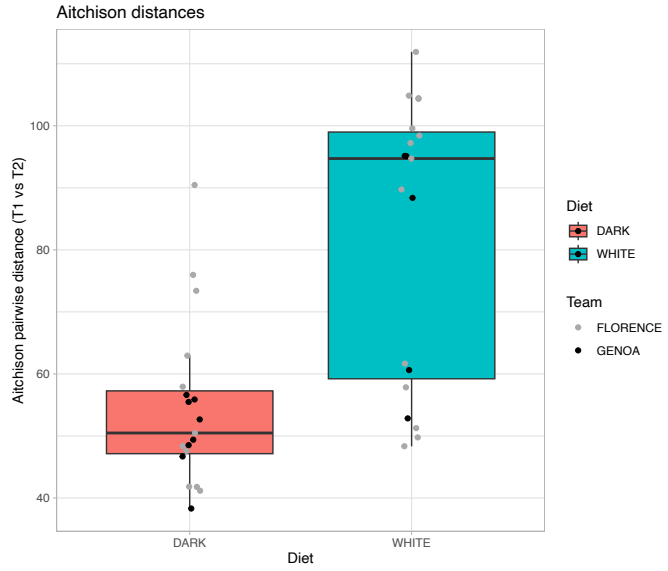


Figure 17. Box plot based on Aitchison dissimilarity distance, wilcoxon non paired, two.sided, p-value = 0.0001226



Microbiome variations are associated with player's name, treatment, team location and ethnicity

Exploring host or environmental factors significantly contributing to interindividual microbiome variation across longitudinal intervention, we combined shotgun metagenomic data with metadata of athletes. Player_name (n=38, Aitchison Distance for β-diversity pre and post intervention, R^2 = 62,3%, P_{adj} = 0,001), treatment (R^2 = 2,3%, P_{adj} = 0,001), time point (R^2 = 1,2%, P_{adj} = 0,003), team_location (R^2 = 2,1%, P_{adj} = 0,001), and ethnicity (R^2 = 4,2%, P_{adj} = 0,001) were identified as cumulative metadata variables over microbiome variations in a multivariate analysis (permutational multivariate analysis of variance, PERMANOVA). Further, to understand the relative effect size of each significant metadata variable (identifying variables with non-redundant explanatory power over longitudinal microbiome community variation), we investigated variables excluding the covariate "player_name", since this covariate represented the variable with the largest effects size explaining the microbial variation. The exclusion of player_name allowed us to demonstrate that team location, treatment and ethnicity accounted respectively for (R^2_{adj} = 3,6%, P_{adj} = 0,002), (R^2_{adj}= 2,8%, P_{adj} = 0,002) and (R^2_{adj}= 1,6%, P_{adj} = 0,002).

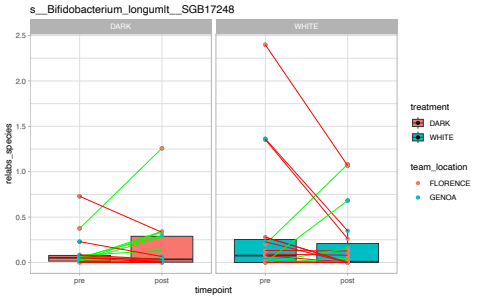
Taxonomic features of the microbiota are associated with dark chocolate intake

To investigate the specific bacterial taxa that were significantly affected by dark chocolate consumption, we compared the composition of the gut microbiota between the dark chocolate group and control group at species level. Although dark chocolate group maintained a greater stability compared to control group, some compositional changes in microbial community have been observed. At species level, Wilcoxon-paired test showed that *Bifidobacterium longum* (n = 38, *two-sided Wilcoxon rank-sum test*, $-\log10(P) > 0{,}28$), *Dysosmobacter welbionis* ($-\log10(P) > 0{,}90$), *Faecalibacterium prausnitzii* ($-\log10(P) > 0{,}61$), *Oscillibacter sp. ER4* ($-\log10(P) > 1{,}08$) slightly increased in dark chocolate group, while *Blautia wexlerae* ($-\log10(P) > -5{,}88$) and *Blautia faecis*

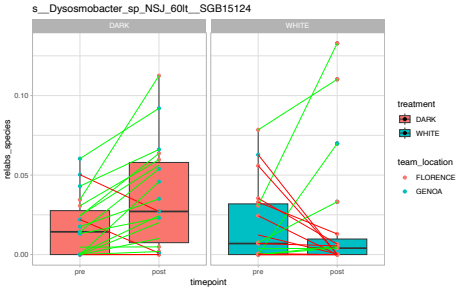
(−log10(P) > -2,27) decreased, compared to the controls. These results identify the effects of dark chocolate consumption on compositional changes in gut microbiota. (Figure 18 a,b,c,d,e).

Figure 18. Compositional changes in microbial communities. Two-sided Wilcoxon rank-sum test.

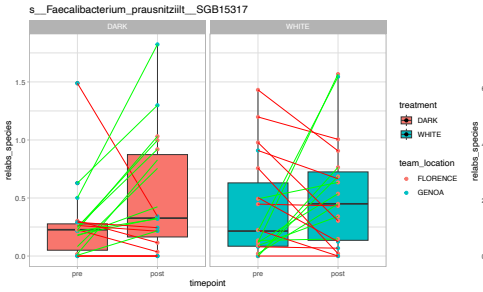
a



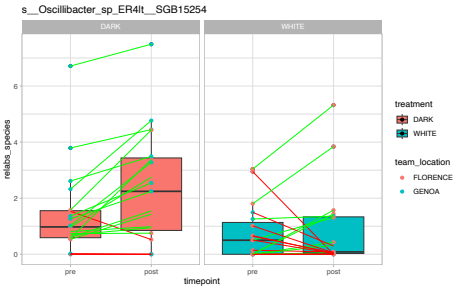
b



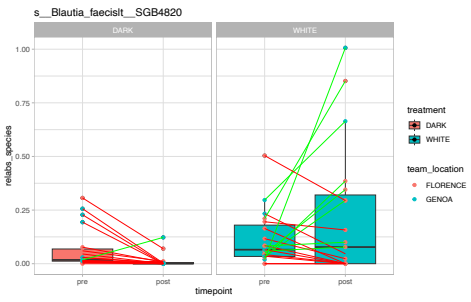
c



d



e



Discussion

Dark chocolate contains a number of polyphenolic compounds, and it is particularly rich in flavonoids - specifically, flavanols, also called flavan-3-ols [70]. Several clinical trials looking at the association between dark chocolate flavanols intake and lipid profile showed that dark chocolate and cocoa rich in flavanols can improve lipid profile status and modulate the metabolism of PUFA in humans [52,

71, 72]. However, to the best of our knowledge, there is no evidence of such effect on elite athlete population, thus, the purpose of this study was to investigate the influence of 30g of dark chocolate on plasma lipid profile, biomarkers of inflammation (AA:EPA ratio), and gut microbiome composition in elite soccer athletes. In this research we extended to elite athletes' population the previous observations that dark chocolate and cocoa rich in flavanols can improve lipid profile and the metabolism of PUFA in humans [52, 71, 72], and additionally, we demonstrated that dark chocolate may contribute to the maintenance of a stable microbiota during intense exercise training. It is known that high total and LDL cholesterol, or low HDL cholesterol represents a significant long-term risk factor for cardiovascular diseases (CVD) [73]. However, individuals who practice sports involving a high level of physical exertion, such as elite soccer players, show a less favorable lipid profile, compared to healthy population, thus potentially increasing the CVD risk factor [74-76]. Although it is difficult to identify the cause of why the most stressful exercise should impair the lipid profile towards a dysmetabolic pattern, it has been speculated that intense exercise, and particularly eccentric exercise accompanied by falls and collision [77], leads to the release of proinflammatory cytokines [78] associated with abnormalities in metabolism.[79, 80]. In addition, exercise-associated perturbations of gastrointestinal epithelial wall integrity, may lead to translocation of pathogenic bacteria and/or bacterial endotoxin and subsequently induce local and systemic inflammation [81] (cytokinemia). Further, it has been shown that lipopolysaccharide (LPS), the major component of the outer membrane of Gram-negative bacteria, can impair lipid metabolism increasing the activity of 3-hydroxymethyl-gluaryl coenzyme A reductase and reducing the level and activity of cholesterol 7α-hydroxylase mRNA [82]. For all these reasons, it seems reasonable to identify potential nutritional strategies aimed to decrease the inflammatory status of elite athlete and, consequently, the risk for CVD. Interestingly, our study revealed that total cholesterol, LDL cholesterol and triglycerides significantly decreased after 4 weeks of dark chocolate supplementation and, in addition, HDL cholesterol level significantly increased in the DC group. Nonetheless, no effects were evident in the control group. Our findings are in line with a recent clinical trial which observed the same positive

effect of cocoa supplementation (~220mg flavanols) on the improvement of lipid profile in a cohort of healthy population [5], and with two meta-analyses demonstrating the beneficial effects of dark chocolate/cocoa products on total and LDL cholesterol [83, 84]. Moreover, our study was also designed to assess whether dark chocolate intake could modulate the metabolism of PUFA by decreasing the plasmatic AA/EPA ratio. We chose the AA/EPA ratio as biomarkers since its measurement may serve as a marker for cardio-protective effect, such as anti-inflammatory effect [85]. Our results showed that dark chocolate may interfere with the metabolism of PUFA by significantly decreasing the AA/EPA ratio in elite soccer athletes. The mechanism by which cocoa flavanols elevate plasma EPA level or decrease AA concentration remain still uninvestigated, however, because of the control group did change the AA/EPA ratio level, it is plausible the hypothesis that the polyphenol content of dark chocolate may be responsible of this effect. Lastly, we looked for the potential effect of dark chocolate on athlete's gut microbiota community. Gut microbiota and exercise have recently been shown to be interconnected [22]. However, if on one side, moderate exercise (< 70% of maximal oxygen uptake, VO2Max) has positive effects in the gut microbiome, on the other, intense exercise (> 70% VO2Max) usually practiced by elite athletes, can increase blood markers of inflammation [86] and impair gastrointestinal epithelial permeability, thus enabling pathogens to enter the bloodstream and increase inflammation level [81, 87]. For these reasons, we decided to investigate whether the daily intake of dark chocolate, known to have prebiotic-like effect [5], may positively affect the gut microbiota of elite soccer players. In the present study, we observed that 30 days of 30g dark chocolate did not affect the overall gut microbiota composition of both DC and WC group in terms of alpha diversity (Faith's PD) and richness indices. However, subject variance analysis, which allowed us to measure the microbial stability within individual participants, showed that the microbial community of each individual in the DC group was more stable over time exhibiting lower 'within-subject community dissimilarity' compared to subjects of control group. Conversely, athletes in the control group exhibited greater microbial turbulence with bigger shifts in the gut communities across the study. There results suggest that dark

chocolate may be linked with gut microbiota where athletes provided of 30g dark chocolate/daily maintained a more stable gut ecosystem compared to controls. Interestingly, the recent body of evidence demonstrates that gut microbial stability and resilience represent essential ecological characteristics of gut microbiota and measures associated with host healthy states [88]. Accordingly, in the case of athlete population, the maintenance of gut microbial resilience after different dietary perturbations was associated with greater athletic performance [89]. Furber et al. [89] recently demonstrated that subjects consistently performed better during athletic performance test when microbial communities remain relatively unchanged throughout dietary challenges (high protein or high carbohydrate diets). Indeed, the maintenance of a subject-specific ecosystem and reduced longitudinal variation was associated with higher Time to Exhaustion (TTE) performance in both dietary interventions. In other words, athletes undergoing dietary periodization [90] or dietary supplementation protocols [91], with the aim of improving performance, would likely benefit from greater gut microbial resilience. Thus, from a clinical and translational perspective, the ability of dark chocolate to positively modulate lipid profile and PUFA metabolism while maintaining gut microbial stability, represents an effective nutritional therapeutical strategy for improving and/or maintaining elite athlete's health. Our result stressed the concept that, although microbiome states that are unequivocally "healthy" are yet to be established [92], "it is less the response to perturbation but rather post-perturbation resilience the greatest hallmark of health" [1]. However, our explorative analyses of gut microbiota variation across 1-month dark chocolate intervention revealed the microbiome of individuals to be fundamentally unique, with player inter-variability explaining the largest effect of microbial community variation across study. In line with these findings, Johnson et al., demonstrated that microbiome composition is highly individual and daily microbial responses to diet is highly personalized [37]. Further, Wu et al., revealed that that inter-subject microbiome variation remained high even after periods of identical dietary intake [93]. Lastly, although dark chocolate group maintained a greater stability compared to control group, significant changes have been detected at species level. The top six flux changes in relative abundance that we observed were

an increase in *Bifidobacterium longum, Dysosmobacter welbionis, Faecalibacterium prausnitzii, Oscillibacter sp. ER4* and a decrease in *Blautia wexlerae* and *Blautia faecis* in the dark chocolate group compared to controls. *Bifidobacterium longum* is known to be a probiotic bacterium with several potential benefit for health promotion such as immune regulation, anti-tumorigenesis, recovery of healthy gut microbiota after antibiotics treatment, and vitamin biosynthesis [1, 94]. Interestingly, *bifidobacteria* genera have been reported to have an anti-cholesterolemic effect. The available evidence suggested that *bifidobacteria* genera reduce blood cholesterol in two different ways: i) bifidobacterial cells absorb and accumulate intestinal cholesterol, and they are gradually excreted with the feces from the intestine [95]; ii) cholesterol is a precursor of conjugated bile acid. *Bifidobacteria* have bile acid hydrolase to convert conjugated bile acid to an unconjugated one. After deconjugation of bile acid, this unconjugated one is precipitated and excreted with the feces from the intestine. To recover the level of bile acid, cholesterol is gradually converted to bile acid in the liver, resulting in cholesterol reduction in the blood[3]. More specifically, Kim et al. showed that *in vivo* administration of *Bifidobacterium longum subsp. BCBL-583* to a high-fat mouse model (HFD-583) solved the undesirable problem regarding obesity caused by the administered HFD. Indeed, *B. longum BCBL-583* reduced the total cholesterol and LDL-cholesterol, decreased pro-inflammatory cytokines and increased anti-inflammatory cytokines, substantiating its cholesterol reduction and anti-inflammation activities [96]. Differently from the recent findings observed by Shin et al. [97], we observed an increased in *Faecalibacterium prausnitzii* species. *F. prausnitzii* is a butyrate producers' bacteria with anti-inflammatory properties and represent a crucial community for the maintenance of gut homeostasis [98]. Moreover, *F. prausnitzii* has been found in greater relative abundance in active women (with an increase associated with exercise training [99]) and in microbiomes of professional athletes participating in sports with a high dynamic component (>70% VO2Max [100]). Therefore, we hypothesized that the combination of intense exercise training and dark chocolate ingestion might have synergically contributed to increase the levels of *F.prausnitzii*. Indeed, in the study of Shin et al.[97], in which the authors observed a decrease in *F.prausnitzii* after dark chocolate

supplementation, the cohort of study was selected among healthy adults with a sedentary lifestyle and it is known that the microbial profile commonly observed in professional athletes differ significantly from sedentary individuals. The intestinal microbiota of athlete is mostly composed of good bacteria, such as *F. prausnitzii*, and is characterized by higher levels of propionate, acetate and butyrate producing bacteria [101]. In line with these metabolic properties, we revealed an increase in *Dysosmobacter welbionis* species. *Dysosmobacter welbionis* has been recently discovered to be a butyrate producing bacteria negatively correlated with body mass index (BMI) [102]. Le Roy et al., showed that the supplementation of *Dysosmobacter welbionis* (strain J115T) to HFD-fed mice reduced white adipose tissue hypertrophy and inflammation together with increased number of mitochondria in brown-adipose tissue and non-shivering thermogenesis. This finding was associated with improved glucose homeostasis and increased non-shivering thermogenesis, thus beneficially influences host metabolism [102]. In addition, we observed an increase in the genus of ***Oscillibacter*** that has been detected as the major determinant of obese or normal status [103]. Indeed, Tims et al. examined twins that were discordant in terms of BMI status and found that *Oscillibacter* genus was more abundant in the leaner twin [104]. Finally, we observed a reduction in *Blautia wexlerae* and *Blautia faecis* species. *Blautia* genus has been shown to be one of the most abundant genera in prediabetes and T2D compared with healthy subjects [105] and has been suggested to increase the release of pro inflammatory cytokines (TNFα, cytokines) [106]. Moreover, Motiani et al. showed that *Blautia* decreased after 2 weeks of training and that the reduction was associated with a decrease in the TNFα. This reduction is important as TNFα plays a critical role in the inflammatory processes, such as IBD [107]. In our study lower abundance of *Blautia wexlerae* and *Blautia faecis* species was observed in DC group, thus suggesting that dark chocolate supplementation may have reduced the inflammation in the gut and at systemic level. However, Shin et al., observed a higher abundance of the genus of *Blautia* in healthy population after 3 weeks of dark chocolate intervention [97]. Based on these conflicting results, it is still necessary to elucidate the characteristics of *Blautia* for specific species (or even strain) and to perform the direct feeding of this bacterium to human participants to

clarify the role of this bacterium for correlation with human health in the future. Although data on the impact of polyphenols on the gut microbiota and their mechanisms of action in humans are scarce [5], we demonstrated that dark chocolate and the gut microbiota may interact and positively influence athlete's health. Moreover, we showed that dark chocolate supplementation contributes to the maintenance of gut health by the modulation of the gut microbial community (stability over time) and through the stimulation of the growth of beneficial bacteria. However, these results raised the questions of whether dark chocolate supplementation directly affected the growth of the *Bifidobacterium longum, Dysosmobacter welbionis, Faecalibacterium prausnitzii* and *Oscillibacter sp. ER4* or if it was a combination of indirect effects via systemic changes. Whatever the nature of the prebiotic effect of dark chocolate, this is the first study conducted to demonstrate that ingestion of a dark chocolate may have these new properties on elite athlete's population. However, we acknowledge that our study has several limitations. For example, we selected our cohort of elite soccer players from two different geographic locations (Genova and Florence, Italy) and it has been demonstrated that team location influenced the response in terms of microbiota changes. Accordingly, recent studies have showed a number of divergences in the microbiome composition between healthy individuals from different geographic locations (cultural/behavioral features) [108]. It should be highlighted that we decided to choose these two teams because of the same number of matches played (same volume of training volume) within the competitive season. However, we adopted appropriate statistical analyses (linear mixed model) to minimize the bias. In addition, another limit is that the underlying mechanism by which dark chocolate modulate the intestinal microbiota and elevate plasma EPA or DHA concentrations (or decrease AA levels) remain unclear. This encourages future studies to examine gut-liver signaling, such as associations between bacterial metabolites and cholesterol, following a period of dark chocolate consumption. Despite its limitations, to the best of our knowledge, our data suggest that supplementation with 30g dark chocolate, over 4 weeks of intense exercise, may improve lipid profile, the metabolism of PUFA and the gut microbiota community in elite soccer players. With the current advances in technology, metabolic profiling, and

others “omics” analysis can be harnessed to discover candidate microbial communities and their derived metabolites that impact human health [109]. Such advances will help discern the gut microbial mechanisms that explain the relationship between intestinal microbiota with cholesterol metabolism. This will lead to deeper understanding of the microbiota and encourage the development of microbiota-based therapies and microbiome-informed precision medicine to reduce the global risk for CVD.

PART 2 - STUDY 3:

Effect of thirty days of ketogenic Mediterranean diet with phytoextracts on athletes’ gut microbiome composition: introduction

See our systematic review for the general background:

*Paoli A, **Mancin L**, Bianco A, Thomas E, Mota JF, Piccini F. Ketogenic Diet and Microbiota: Friends or Enemies? Genes (Basel). 2019 Jul 15;10(7):534. doi: 10.3390/genes10070534. PMID: 31311141; PMCID: PMC6678592* for further information.

The human intestinal tract is composed of a considerable population of microorganisms (microbiota) and its corresponding gene complement (microbiome), that symbiotically live within the host. In recent years, the awareness of the importance of microbial community in human health has increased tremendously, making the science of microbiome a key area for life sciences [110]. Intrinsic and extrinsic factors including age, environment, birth delivery route, breastfeeding, antibiotics, genetic background, human leukocyte antigen, dietary factors, and exercise, impact the microbial composition and function, with the diet and exercise act as primary modulators [67, 111-115]. More specifically, in sport nutrition, diet represents one of the most important tools that athletes use to optimize their fitness, performance and recovery and macro nutrients manipulation are often adopted to optimize training outcomes and competitions’ performance. For example, carbohydrates represent

a primary fuel source during physical activity, and they are fundamental to maintain and refill athlete's muscle glycogen stores. To date, recent evidence suggests that carbohydrates may influence athletic performance also *via* the modulation of gut microbiome [116]. Indeed, the effect of carbohydrates on the gut microbiome differs widely as a function of microbiota-accessible carbohydrates (MACs) commonly referred to as dietary fiber, content, and types. Dietary MACs are found in a variety of sources including plants, animal tissue, or food-borne microbes and represent the source of carbohydrates that are metabolically available for gut microbes. MACs hold a role of "primary fermenters" within the colonic ecosystem and generally tend to increase the production of the beneficial short chain fatty acids (SCFAs) producing bacteria such as *Bacteroides*, *Firmicutes* and *Actinobacteria* [117]. Differently, an increased consumption of protein among athletes, may lead to an excessive protein fermentation associated with the increased abundance of related taxa such as *Clostridium* and *Proteobacteria*. In sport nutrition, an additional area of interest is also represented by the study of ketogenic diet (KD) effects on athletes' health and performance. Indeed, high fat-low carbohydrate diet, such as ketogenic diet, has gained popularity among athletes and practitioners for its potential application in sports [118]. KD represents a dietary protocol consisting of high-fat, adequate protein and less than 20g of carbohydrate daily (or 5% of total daily energy)[118]. This nutritional approach has been used since the 1920 as a treatment for refractory epilepsy [119] and it has gained popularity as a potential treatment for obesity and related metabolic disorders [120]. Indeed, increased amount of evidence point out that KD may represent an efficient and safe solution to get adequate body composition and maintain a general good health. The metabolic shift induced by ketogenic diet and some of the complex metabolic pathways involved in 'ketotic state' has suggested a possible use of ketogenic diet in sports[118]. For example, the use of KD may represent a safe strategy for the athletes who need to reduce body weight and body fat while maintaining lean mass and performance [121]. One of the concerns raised about the use of KD for sport purposes is related to its putative negative impact on gut microbiome [122]. On the other side, substantial changes in microbiome composition have been also attributed to exercise. To date, some studies reveal that

exercise may increase the gut microbiota diversity and associated microbial-derived metabolites [45, 67]. Observational studies have revealed that high-level athletes have an increased microbial α-diversity (a measure of microbiome diversity of a single sample), lower inflammatory markers and a higher microbial production of short chain fatty acids (SCFAs) [45]. For example, Clarke et al., compared the gut microbiota of professional Irish male rugby players with two groups of healthy, non-athletes subjects matched for body mass index (BMI): (> 28kg/m2) and (< 25kg/m2) and found that the microbial diversity of rugby players was higher compared with both non-athletes groups [67]. More recently, Scheinman and colleagues collected and sequenced the stool samples from a cohort of athletes participating to the Boston Marathon (1 week before and 1 week after), along with a group of healthy- non athletes' controls. The researchers found that the most differentially abundant specie was *Veillonella atypica*, a Gram-negative bacterium that metabolize lactate into acetate and propionate via the methylmalonyl-CoA pathway. Further, compared with mice gavaged with *Lactobacillus*, the transplantation of stool containing the *Veillonella* significantly improved submaximal treadmill run time to exhaustion, suggesting a potential role for *Veillonella atypica* in improving athletic performance. The authors suggested the possibility that the lactate produced during sustained exercise could be converted by *Veillonella atypica* into propionate, identifying a new microbiota-driven enzymatic process that may improve athletic performance [112]. To the best of our knowledge, only one study investigated the effect of KD on the gut microbiota in athletes (a cohort of elite race walkers) [123], while no studies are available in a model of mixed endurance/power sport such as soccer. In our recent article [121] we investigated the effect of 30 days of ketogenic diet on body composition, muscle strength, muscle area and metabolism in a cohort of semi-professional soccer players. The athletes who underwent the KD intervention lost body fat mass without detrimental effect on strength, muscle mass and power. However, considered the suggested detrimental effect of KD on gut microbiome [123], the aim of the current study was to assess the gut microbiome composition of semi-professional soccer players who participated in the above cited

study, to understand whether and how the gut microbiota changes in response to thirty-days of ketogenic Mediterranean diet with phytoextracts (KEMEPHY) diet.

Material and methods

Participants

This is a secondary analysis of a previous published research [121]. A more detailed description of the experimental study and physiological measures can be found [121]. Sixteen semi-professional soccer players (25.5 ± 2.8 years, 77.2 ± 11.88 kg) were recruited for the study. The exclusion criteria were: participants with a body fat percentage over 32%, (determined *via* dual energy X-ray absorptiometry DXA), cardiovascular, respiratory, gastrointestinal, thyroid or any other metabolic diseases, weight change ± 2 Kg over the last month, adherence to special diets, use of nutritional supplements (except a daily multivitamin-mineral), use of antibiotics [124], use of medication to control blood lipids or glucose. The anthropometric details of the subjects enrolled in the study were provided in our previous published study [121]. During the study players were asked to keep their normal training schedule (8 hours of training/week). After the medical health screening, all the subjects read and signed the informed consent with the description of the testing procedures approved by the ethical committee of the Department of Biomedical Sciences, University of Padova, and conformed to standards for the use of human subjects in research as outlined in the Declaration of Helsinki, Clinical Trial registration number NCT04078971.

Study design and procedures

The study was a randomized, parallel arm, controlled, prospective study in which gut microbiota was tested before and after thirty days of KEMEPHY protocol. Subjects undergone to several anthropometric and performance measurements described in our previous paper[121]. Subject were randomly assigned to the KEMEPHY diet (KDP n=8) group or Western Diet (WD n=8) group,

through an on-line random number calculator (https://www.graphpad.com/quickcalcs/randMenu/), matched for percentage of body fat. The workload of all athletes was over-imposable because the coach and trainers strictly controlled the training schedule, and they were instructed to maintain the same level of physical activity throughout the study. (The study protocol is shown in detail in our previous article).

Dietary intervention

Before the start of the study, athletes were provided nutritional counseling and resources to better adhere to KEMEPHY. Resources included food lists containing the food prohibited and permitted in ketogenic diet and electronic-suggested daily meal plans, meal recipes. The food lists encouraged on eating unprocessed meat including beef, veal, poultry; fish such as eel, mackerel, salmon, sardines; raw and cooked vegetables, cold cuts such as dried beef, eggs and seasoned cheese (parmesan); Konjac; fruits with the lowest glycemic index (blueberry, raspberry), raw nuts and seeds, ghee butter, butter, plant oils and fats from avocado, coconut and green olives [125]. The drinks permitted were tea, coffee, herbal extracts without sugar and it was allowed a "*Keto cocktail*" once a week, made up of gin and soda. Moreover, since the nutritional protocol of KD it may be hard to be maintained for long periods due to the lack of sweet taste[126], many ready-to-eat ketogenic products (RKP) have been provided in addition to usual low carbohydrate foods [127]. The present study indeed tested some ready-to-eat foods selected from the product range of Tisanoreica® snacks and meals (Gianluca Mech S.p.A., Asigliano Veneto, Vicenza, Italy) and Le Gamberi Food ® and meals. In our protocol we used some RKP as a ketogenic pasta (selected with a ketogenic ratio of fats: protein+carbohydrate equal to 4:1) (Le Gamberi Foods, Forlì, Italy), and other RKP (specialty meals and drinks) that mimics the taste of carbohydrates, constituted principally of high-quality protein (18 grams of protein per portion), fibers, and electrolytes (mainly magnesium and potassium) (Tisanoreica® by Gianluca Mech S.p.A., Asigliano Veneto, Vicenza, Italy), detailed in Table 1. Among the products selected, there were 4 sweets RKP products (chocolate biscuits CB (Cioco-Mech); chocolate and hazelnut balls

CHB (Bon Mech); apple-cinnamon biscuits ACB (T-Biscuit); chocolate-almonds-pistachio bar CAPB (T-Smart) and one savory product: pasta P1 (Le Gamberi Pasta). Both diets were designed to be isoproteic i.e., same amount of protein (1.8g x Kg^{-1} x body $weight^{-1}$ x day^{-1}). The distribution of macronutrients during the KEMEPHY was carbohydrate (<30gram x day^{-1}; <10%) protein 1.8g x Kg^{-1} x body $weight^{-1}$ x day^{-1} (~25-30%), fats *ad libitum.* Moreover, each subject was provided of three herbal extracts (Paoli et al., 2021) according to commercial ketogenic protocol (Tisanoreica®, Gianluca Mech S.p.A., Asigliano Veneto, Vicenza, Italy). During the first week, subjects were provided of pure medium chain triglyceride oil (MCT oil: 20g Named® Natural Medicine), in order to facilitate ketosis [128] and to allow players maintaining the same work load during training sessions. WD group was provided of a diet similar to western diet, thus the intake of protein has been increased to 1.8g x Kg^{-1} x body $weight^{-1}$ x day^{-1} in order to be make the two diets isoproteic. The WD was composed mainly of whole cereals (spelt, rye, oat) and pseudo-cereals (buckwheat, quinoa, amaranth), whole grain pasta, potatoes, meet, fish, vegetables, fruit, legumes, olive oil, milk, and red wine (at most 1 glass per day). Thus, the WD ensured a constant energy and macronutrient balance: protein 1.8g x Kg^{-1} x body $weight^{-1}$ x day^{-1}, (~ 30%), fats ~20-25% and carbohydrate ~50-55%. WD diet was also designed to contain <10% saturated fat and <300mg cholesterol/day. It should be stressed that, as it can be noted, the WD diet we provided to the athletes was totally different from the typical high-fat, high sucrose Western diet usually adopted in research studies. In both groups protein intake was distributed equally throughout the day (every 3 - 4 hours) and pre-sleep casein protein intake (30 - 40 g) was provided in both group after training evening session, as indicated by the ISSN's position stand [129]. The diets were explained to all subjects during an individual visit and dietary intake was measured by validated 3-food-diary that has been used in the past in studies with athletes [130] and analyzed by *Nutritionist Pro™ (AxxyA systems, Arlington, VA).* Subjects received the specific instruction for completing detailed weighed food records during 7 day-periods for each diet and were daily monitored by call interviews each day after dinner. To ensure that carbohydrates were restricted throughout the KEMEPHY diet, subjects tested their urine daily using

reagent strips at the same time of the day (Ketostix semiquantitative urine strips, Bayer, Leverkusen, Germany), recording the result on log sheet and, once or twice a week, subjects were tested by *GlucoMen LX Plus (Menarini Diagnostics, Firenze, Italy)* to detect ketones concentration in capillary blood. Subjects received follow-up counselling and dietetic education if necessary. Additionally, a *WhatsApp* (Meta Inc., Mountain View, CA, USA) group was created and some applications for smartphone were provided (*Keto-diet tracker, https://ke.to*; Keto-app*, https://ketodietapp.com),* to track their food daily intake.

Faeces sampling and DNA extraction

Faeces samples were collected at baseline and after thirty days of dietary protocol. 100-150mg of faeces were collected using sterile swab (FLmedical, Italy) tubes (Starlab Group, Italy) and preservative buffer (Zymo Research, USA) in the morning of the day of starting KEMEPHY and after thirthy days. Samples were sent to BMR Genomics srl (Via Redipuglia, 22, 35131 Padova, PD) within 2 days and stored at −20 °C until DNA extraction. DNA was extracted using Cador Pathogen 96 QIAcube HT Kit (Qiagen srl, DE) with lysis step modification according to Mobio PowerFecal kit (Qiagen srl, DE).

16S rRNA gene sequence data processing and analysis

The V3-V4 regions of the 16S ribosomal RNA gene were amplified using Illumina tailed primers Pro341F (5′-TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG-CCTACGGGAGGCAGCA-3′) and Pro805R (5′-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGACTACNVGGGTATCTAATCC-3′) using Platinum Taq (Thermo Fisher Scientific Inc, USA) by means PCR (94° C for 1 min, followed

by 25 cycles at 94° C for 30s, 55° C for 30s, and 68° C for 45s, and a final extension at 68°C for 7 min). PCR amplicons were purified by means Agencourt AMPure XP Beads 0.8X (Beckman Coulter, Inc., CA, USA) and amplified following the Nextera XT Index protocol (Illumina, Inc., CA, USA). The indexed amplicons were normalized by SequalPrepTM Normalization Plate Kit (Thermo Fisher Scientific Inc.) and multiplexed. The pool was purified with 1X Magnetic Beads Agencourt XP (Beckman Coulter, Inc.), loaded on the MiSeq System (Illumina, Inc.) and sequenced following the V3 - 300PE strategy. The bioinformatic analysis was performed by means QIIME 2 2021.4 version [131]. Raw reads were firstly trimmed applying Cutadapt to remove residual primer sequences and then processed with DADA2 plug-in [132] to perform the denoising step. DADA2 was run with default parameters except for the truncation length: forward and reverse reads were truncated at 260 and 245 nucleotides, respectively. The resulting Amplicon Sequence Variant (ASV) sequences were filtered out by applying a 0.01% frequency threshold in order to discard singletons and very rare sequences. All the samples included in the analysis was rarefied. The value of rarefaction is 32232 reads (Supplemental Figure 1). The more recent available Silva 138 database [133] as used to associate the taxonomy to the remaining ASVs for the final analysis; moreover we earlier performed also an analysis with Green genes v.13-8 database that will be briefly discussed to better understand the variability due to the database utilized.

Statistical analysis

Results are presented as mean and standard deviation (SD), or median and quartiles (Q1-Q3) where appropriate. Alpha diversity indexes (OTUs number and Shannon's Effective Number of Species) were computed with the diversity function of the *vegan* R package, and time, group and time×group effects were tested using a Wilcoxon test for paired data (interaction effect was checked while performing the test on delta values); a false discover rate (FDR) with Benjamini-Hochberg correction was applied to account for multiple testing. Effect sizes were calculated with the *rstatix* and *coin* R packages. Common interpretations of Wilcoxon effect sizes (r) are: 0.10 - 0.3 (small effect), 0.30 -

0.5 (moderate effect) and ≥ 0.5 (large effect). A dissimilarity matrix with Bray-Curtis distance was calculated, and a Permutational Analysis of Variance (PERMANOVA) for repeated measures was used to test pre-post differences between the two groups (KDP vs WD) in the relative abundances at phylum and genera taxonomic levels, using the *adonis* R function, and post-hoc comparisons were performed with a paired Wilcoxon test with FDR correction. Furthermore, after ruling out baseline differences in the microbial composition at baseline, data were filtered for the presence of each taxon in at least 70% of the subjects, and a linear discriminant analysis (LDA) was performed at the different taxonomic levels (from phylum to genus) on the post-intervention data (LEfSe; LDA Score > 2.0, $p<0.05$); significant different taxa were graphically represented on a cladogram. To assess correlations between macronutrient intake (7-days food diary) and pre-post treatment variations in body composition, fitness measures and genera abundances, a Spearman correlation matrix was computed: significant correlations were extracted ((Spearman $r_{0.05, 14} \geq 0.503$), and represented in a circular plot using the *circlize* R package. Analyses were performed using R Studio 4.1.1; the significance level was fixed at the standard value of 0.05.

Results

Dietary nutrition intake

There were no differences in dietary nutrient intakes between groups at baseline. Subjects adhered to the given instructions for both diet interventions according to analysis of diets records (3 days food-diary before the study and 7 days food-diary during the study). During the diet interventions, all dietary nutrients were significantly different between the KEMEPHY and WD diets. Indeed, the intake of CHO g/day and % in KEMEPHY and WD group was respectively (KDP = 22 ± 5 g/day; WD = 220 ± 56 g/day, $p<0.0001$), (KDP = 9 ± 3 %; WD = 51 ± %, $p<0.0001$) while the intake of % fat was (KDP = 64 ± 3%; WD = 20 ± 8 %; $p<0.0001$). In addition, the total energy intake was reduced during both the treatments but without a significant difference between groups (KDP = 1.984 ±

340Kcal/day; WD = 1.752 ± 320Kcal/day), ($p>0.05$). The complete results about dietary nutrition intake during the intervention are shown in Table 3 of the previous study [121]. For an easier understanding we reported pre- and post- daily dietary energy and nutrient intake in brief in Table 2.

Table 3. Daily dietary energy and nutrient intake at baseline and during KEMEPHY diet (KDP) and Western Diet (WD).

	KDP Pre	**KDP Post**	**WD Pre**	**WD Post**	**Time*Diet effect (p)**
Total (Kcal/die)	2356±450	1984±340	2146±230	1752±320	n.s.
Carbohydrates (g/die)	350±66	22±5	363±34	220±56	p<0.05
Protein (g/die)	105±20	130±25	121±23	129±28	n.s.
Fat (g/die)	107±20	132±27	110±16	38±10	n.s
Carbohydrates (%)	49±6	9±3	51±4	51±4	p<0.05
Protein (%)	15±3	28±4	14±6	28±3	n.s.
Fat (%)	35±4	64±3	33±2	20±8	p<0.05
Protein (g/Kg bw)	1.37±0.5	1.85±0.3	1.59±0.4	1.83±0.2	n.s.
Saturated Fat (g)	35±10	45±12	36±4	15±3	p<0.05
Monounsaturated Fat (g)	28±6	49±16	27±5	9±5	p<0.05
Polyunsaturated Fat (g)	16±3	21±5	16±9	5±2	p<0.05
Cholesterol (mg)	304±101	720±187	303±98	167±65	p<0.05
Fiber (g)	13±2	10±3	11±9	15±4	n.s

Values are mean ± SD, Analysis performed on 3 days of diet records during habitual diet and 7 days during KDP and WD.

Microbiota composition

As alpha diversity measures, the OTUs number and the Shannon's Effective Number of Species (ENS) were calculated. No significant effects of time ($p=0.056$, ES=0.486 and $p=0.129$, ES=0.388,

respectively for OTUs number and Shannon's ENS), group (p=0.317, ES=0.180 and p=0.809, ES=0.047) or time×group (p=0.999, ES=0.01 and p=0.230, ES=0.315) were found (Figure 19).

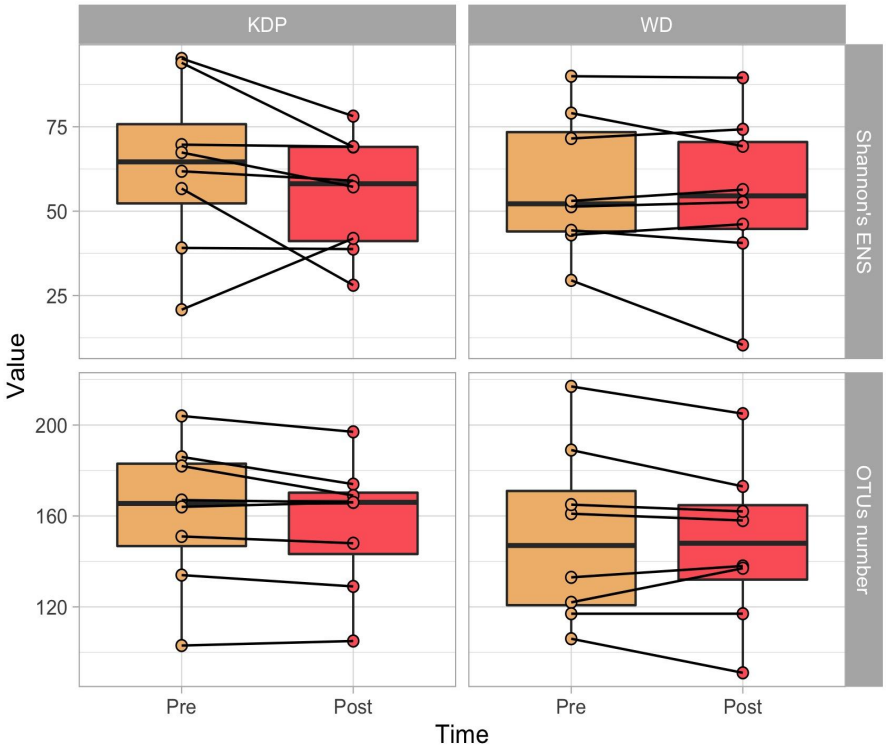


Figure 19. Paired boxplots of OTU's number and Shannon's Effective Number of Species (ENS) in the two groups (KDP vs WD), at the two time points (Pre and Post Intervention).

PERMANOVA for paired data did not find any significant time×group interaction effect for none of the analyzed taxonomic levels (p>0.05). Nonetheless, post-hoc paired Wilcoxon test showed a significant time×group effect for *Actinobacteriota* (p=0.021, ES=0.578), which increased in the WD group (median pre: 1.7%; median post: 2.3%) and decreased in the KDP group (median pre: 4.3%; median post: 1.7%) (Figure 20).

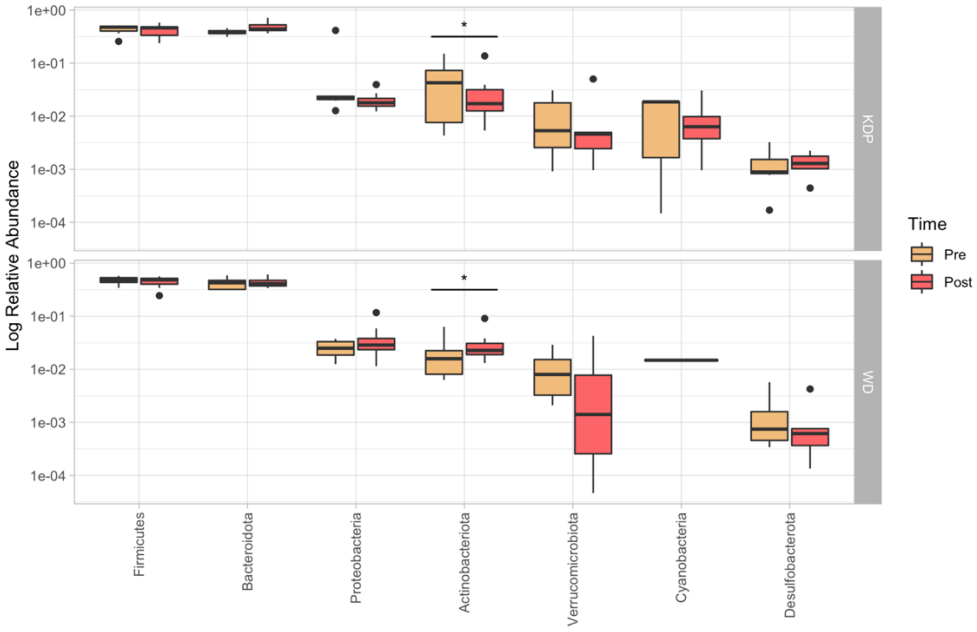


Figure 20**.** Relative abundance (in $\log_{10}$ scale) of the more represented phyla (>0.1%) in the pre- and post-intervention, for KDP and WD groups. Stars represent a significant time×group interaction ($p<0.05$).

Firmicutes/Bacteroidetes ratio was 1.11 (1.07-1.23) in pre and 0.99 (0.73-1.15) in post, and 1.07 (0.99-1.67) in pre and 1.16 (0.94-1.23) in post conditions, in KDP and WD groups, respectively. No significant effect was found for the time×group interaction ($p>0.05$). The linear discriminant analysis in the post intervention differentiated the two groups for *Bifidobacterium* genus (pertaining to the *Actinobacteria* phylum), *Butyricicoccus* and *Acidaminococcus* genera, all more abundant in the WD group, and for *Clostridia UCG-014* (order, family, and genus), *Butyricimonas* and *Odoribacterter* genera (pertaining to the *Marinifilaceae* family), and *Ruminococcus* genus, all more abundant in the KDP group (Figure 21).

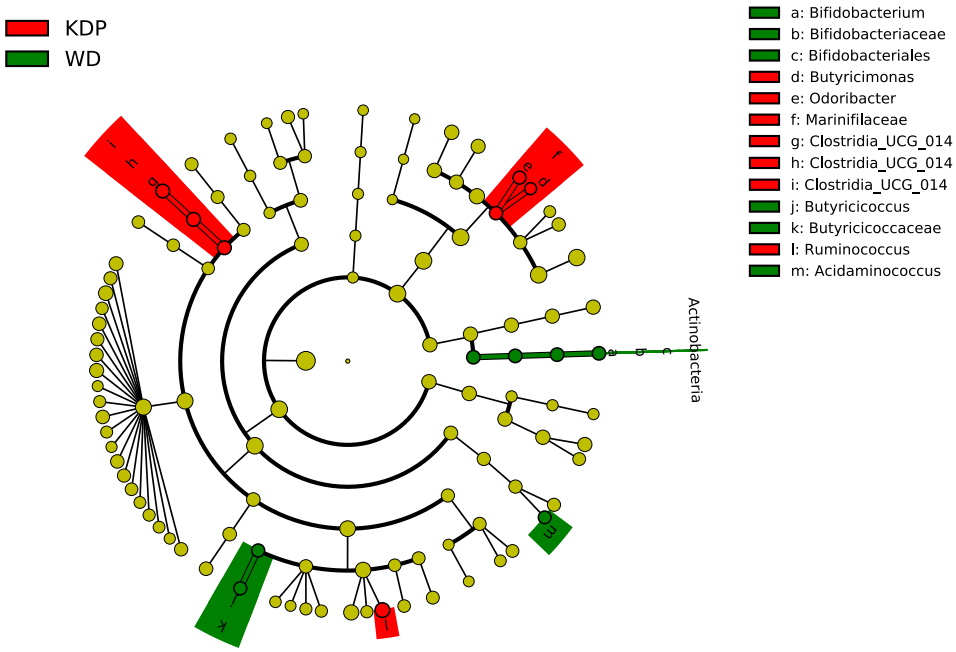


Figure 21. Differential taxa between the KDP and WD groups in the post-intervention (LEfSe analysis, adjusted $p<0.05$, log2 fold change >2).

To investigate the associations between the macronutrient's intake during the intervention and the variations in genera abundances and environmental variables (i.e., anthropometric and performance measures), genera were filtered taking into consideration only those which were present in at least 70% of the subjects, both in pre- and post-interventions. Spearman's correlations were then calculated, and after applying a filter to those statistically significant ($r_{0.05,\,14} \geq 0.503$), were reported on a circle plot (Figure 22). In Figure 22, blue color represents positive correlations while red represents negative ones; the color intensity represents the strength of the correlation. Carbohydrate intake was strongly ($r = 0.84$) associated with a modification in the respiratory exchange ratio (RER), which showed a significant reduction of RER in the KDP group. In other words, players in the KDP group that had less carbohydrate in their diet showed a greater decrease in RER, a sign of an increased reliance on oxidative metabolism. In addition, carbohydrate intake was inversely correlated with changes of *Odoribacter* genus abundance ($r = -0.59$), the latter being also negatively associated to

changes in RER (r = -0.57). This association is coherent with the significant time×group effect in RER presented in Table 3, as *Odoribacter* genus were found to be more abundant in the KDP group (Figure 3). Fat intake, in contrast, was negatively associated with variations of RER (r = -0.68), visceral adipose tissue (VAT) (r = -0.69), extracellular water (ECW) (r = -0.55) and *Fusicatenibacter* genus (r = -0.53). Reductions in weight were associated with a reduced abundance of *Ruminococcus torques* (r = 0.68) and *Lachnospira* (r = 0.71) genera, and inversely correlated with *Parabacteroides* genus abundance (r = -0.62).

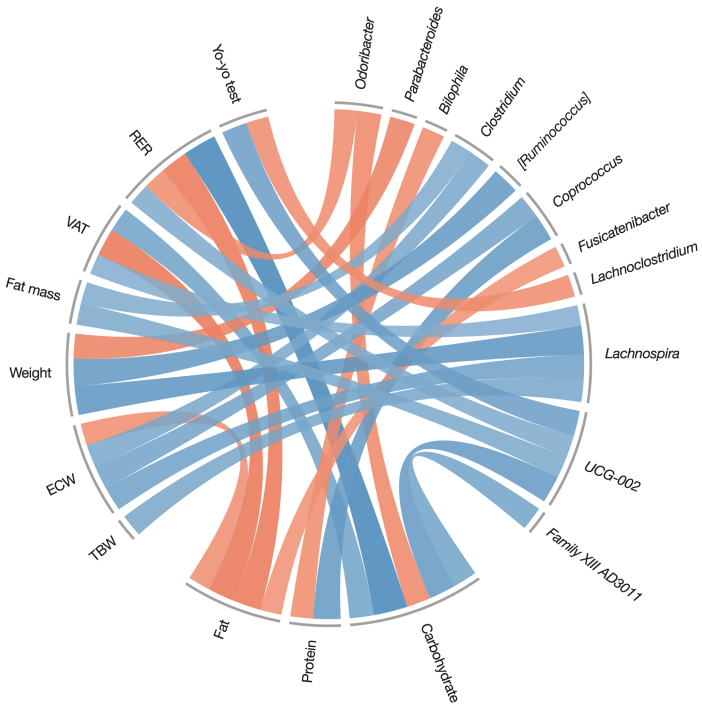


Figure 22. Spearman's correlations between macronutrient intake during the treatment period (7 days food-diary), and post-pre variations on body composition measures, fitness measures, and genera relative abundances. Only significant correlations were reported ($r_{0.05,14} \geq 0.503$). Positive correlations are represented by blue color and negative correlations by red color. Note: TBW: total body water; ECW: extracellular water; VAT: visceral adipose tissue; RER: respiratory exchange ratio.

Discussion

The human gut microbiome is well recognized to be implicated in the promotion-maintenance of health as well in some disease states [134]. Given its plasticity, the gut microbial community can be affected by several factors including genetics, nutrition, environment, exercise and exposure to antibiotics; however, among these contributors, diet elicits the predominant influencing factor [135]. To date, while only one study investigated the effect of ketogenic diet in sport's performance and gut microbiome in endurance discipline [123], no data are available about the effect of ketogenic diet on gut microbiome composition and athlete's performance in team sport. In this study we demonstrate that 30 days of KEMEPHY did not affect the overall gut microbiome of athletes in terms of alpha-diversity indices (the total number of species and the Shannon's Effective Number of Species); however, both groups presented a significant variation both at phylum and genus levels composition (Figure 1). Indeed, the phylum of *Actinobacteria* was significantly decreased in the KEMEPHY and increased in the WD group (Figure 3), while *Clostridia UCG-014*, *Butyricimonas*, *Odoribacterter and Ruminococcus* genera were significantly increased after KDP intervention (Figure 4). Although our data are in contrast with previous studies identifying a positive association between 'high fat diet' and impairment on gut microbiome [14, 34, 35], our results are not surprising since the previous studies investigated the effect of a high-fat, high sugar, Western diet on gut microbiome and did not investigate the effect of ketogenic diet [14, 34, 36] that represent a unique, specific dietary pattern.

In addition, many studies [136-139] investigating the effect of a high-fat diet on gut microbiome tested only mouse models fed a refined high-fat, low fiber diet with animals fed a standard chow diet, high in soluble fibers. For this reason, the conclusions arising from animal studies cannot be adopted to predict the outcomes of a ketogenic diet and, consequently, its associated effect on human gut microbiome [140].

As a matter of fact, in humans, Turnbaugh and colleagues recently confirmed [141] that ketogenic diets differentially alter the composition of gut microbiome when compared to high-fat diet and, further, the authors showed that only ketogenic diet was able to provide positive gut-associated

systemic outcomes [141]. Moreover, another explanation for the maintenance of microbial diversity after KEMEPHY intervention may rely on the specific composition of our KEMEPHY diet. Indeed, when investigating the effect of a ketogenic diet on gut microbiome and health parameters, it should be considered not only the amount of fat (i.e., 70-80% fat from total daily calories), but also the different type and quality of fats. Different types of fat are associated with different effects on the gut microbiome and, consequently, with different effects on intestinal and systemic health [4, 142, 143]. If on one side saturated fats are associated with decreased microbiome diversity [122] in humans, polyunsaturated fat such as omega-3 did not affect microbial diversity and richness. Polyunsaturated fats have the capacity to improve gut epithelial integrity and gastrointestinal health through their ability to produce SCFAs [144]. In our study, the KEMEPHY diet was highly composed in mono-polyunsaturated fat (49±16 g and 21±5 g, respectively) differently from the WD diet which was lower (9±5 g and 5±2 respectively) [121]. We hypothesized that sources of omega-3 fatty acids may have act synergically with ketone bodies to promote an anti-inflammatory state [145], also influencing the intestinal microbiome by increasing the production of SCFAs [4]. However, further studies investigating the hypothesized mechanisms are warranted. Of note, more recently, Furber and colleagues [89] investigated the relationships between gut microbial communities and athletic performance in a cohort of highly trained individuals underwent dietary periodization (high-carbs versus high-protein diet). Interestingly, apart from the taxonomic differences between two dietary interventions, the authors revealed that that better athletic performance was linked with gut microbial stasis, where athletes harboring stable microbial communities consistently performed best in each dietary intervention compared to those with a more turbulent gut microbiome. This result brings to light a pivotal concept: the maintenance of a stable gut microbiome during dietary intervention represents a marker for gut-health and athletic performance [89].

Differences at phylum level

At phylum level, the decrease in Actinobacteria relative abundance could mainly be attributed to a decrease of the relative abundance of the genus *Bifidobacterium* (Figure 3). *Bifidobacteria* are common to the healthy human gastrointestinal tract and represent one of the first colonizers of the mammalian gut. *Bifidobacteria* metabolize complex carbohydrates given that the genome of these bacteria harbors many genes involved in carbohydrate metabolism [146, 147]. The metagenome includes a variety of genes encoding for a specific hexose fermentation pathway, the fructose-6-phosphate [148], which represent the principal pathway for the energy output produced, compared to classical pathways used by other fermentative intestinal bacteria. Indeed, it provides a growth advantage for *bifidobacteria* in the presence of complex carbohydrates [148]. These facts may explain the concomitant proportional decrease of *bifidobacteria* and genes involved in carbohydrate metabolism during KEMEPHY intervention. Accordingly to the reduction in *Bifidobacterium* genus, Turnbaugh and colleagues [141] recently demonstrated in a cohort of over-weight humans that the drop in bifidobacterial genera was correlated with the increase of ketone bodies and positively associated with a decreased intestinal Th17 cell levels and adipose tissues. Given the links between obesity and chronic low-grade inflammation [149], the authors suggested that decreased levels of pro-inflammatory Th17 cells in both gut and adipose tissues during ketogenic diet may be a potential mechanism contributing to the greater efficacy of ketogenic diet in improving some aspects of metabolic syndrome such as glycemic control [150] and reduction in body fat [151]. A decline in *bifidobacteria* has been also observed in weight loss intervention on a macro nutritionally balanced diet, gluten-free diet and low-gluten intervention diet [152, 153], thus, the reduction of *Bifidobacterium* abundance after KEMEPHY intervention may be also attributed to the low intake of cereal grains. On the other side, the higher abundance of *Actinobacteria* phylum after WD intervention may be, at least in part, the consequences of the different amount of fibers given that the intake of fibers decreased in the ketogenic diet (from 13g to 11g per day) while increased in WD diet (from 11g to 15g per day), which could be a strong driver of *Actinobacteria* abundance [154]. Finally, at phylum level, our analysis also revealed that KEMEPHY intervention altered the composition

of the gut microbiome by increasing *Bacteroidetes* and lowering the *Firmicutes* phylum (decreased F:B ratio), compared to WD controls. Even though the F/B ratio is outdated [155], many studies [7, 156-158] have reported that the balance of *Bacteroidetes* and *Firmicutes* may represent an important biomarker for obesity and an indicator of health. More specifically, an increased F:B ratio is commonly associated with dysbiosis, obesity and negative metabolic outcomes [159]. These findings are in line with our results since athletes following KEMEPHY underwent a significant reduction in body weight, body fat mass, waist circumference and visceral adipose tissues [121]. Moreover, it is well known that an excess of adipose tissue (and particularly visceral adipose tissue, VAT) is related to inflammation [160]. In our study, both groups lost body weight, but KEMEPHY group showed a greater reduction of fat mass and VAT.

Differences and genus level

At genus level, we observed an increased in *Butyricimonas, Clostridia UCG_14, Odoribacter* and *Ruminococcus*. Enrichment of *Butyricimonas* negatively correlated with BMI and triglyceride levels indicates that these taxa may promote health or contribute to the prevention of obesity [161, 162]. Our results may support this idea because these taxa increased after KEMEPHY intervention. Moreover, a high abundance of butyric-acid-producing such as *Butyricimonas* has been associated with normal weight and diets high in animal protein and saturated fats [163]. Differently as expected, we observed an increase in the relative abundance of the *Ruminococcus* genus in the KEMEPHY group. This result is in contrast with previously data which reported an inverse association between *Ruminococcus* abundance and a poly-unsaturated fat-rich diet [164]. Indeed, the growth of the genus *Ruminococcus spp.* is usually supported by dietary polysaccharides [111] and individuals consuming animal-based diet or ketogenic diet tend to decrease the levels of the butyrate-producing *Ruminococcus spp.* which are mainly involved in the metabolization of undigested complex dietary carbohydrates and production of SCFAs [111]. However, we may speculate that the daily intake of fiber (cellulose, pectin and lignin) provided during KEMEPHY intervention in the food form of

fermented foods, berries and vegetables, was adequate to support the growth of *Ruminococcous* bacterial taxa. Accordingly, we also observed that *Odoribacter* genus increased after KEMEPHY intervention. *Odoribacter*, belonging to the order *Bacteroidales*, is a common SCFAs producing bacteria [165], and, it seems to be associated with some metabolic health benefit such as the improvement of obesity condition [166, 167].

Importance of up-to-date database

To underline the importance to utilize an up-to-date database in such a new and rapidly growing field as microbiome analysis we reported here, briefly, the most significant differences between our previous analysis performed with Green genes v.13-8 database and the current Silva 138 database. The almost daily advancement in new bacteria classification request the utilization of the most recent database Silva 138. To confirm this fact, the same data showing the main differences are presented in brief and showed in Supplemental Table 1 and Supplemental Figure 2.

Green genes v.13-8 database vs Silva 138 database

At phylum level the differences in *Proteobacteria* disappeared with the more recent database, while the phylum of *Actinobacteriota* did not change: it increased in the WD group and decreased in the KEMEPHY group. At genus level, the main differences were found for *Ruminococcus* and *Dorea* genera. In the previous analysis both genera were slightly reduced in the post condition for KEMEPHY and increased in the WD group, while, with the recent Silva 138 database, the genus of *Ruminococcus* increased in KEMEPHY group while *Dorea* disappeared. More specifically, Green gene database revealed an increase in *Bifidobacterium*, *Roseburia, Butyricicoccus* and *Gemmiger* genera in the WD group, and an increase in *Parabacteroides* and *Odoribacterter* genera for KEMEPHY group; differently, the last database revealed an increase in Clostridia UCG-014, *Butyricimonas* and *Odoribacterter* genera in the KEMEPHY group, while the genus of *Paracteroides* disappeared.

The potential mechanisms of positive effects of KEMEPHY diet on gut microbiome

Our findings suggest that ketogenic diet may partially affect the intestinal ecosystem throughout different mechanisms. We hypothesized that one of these mechanisms might include the production of SCFAs and especially butyrate. Indeed, we supposed that during ketogenic diet, SCFAs and butyrate may be originated from:

i) the liver and then secreted into the gut (because of the ketogenic state);

ii) ketogenic regimens adequately formulated for supplying a medium but adjusted amount of plant-based fermentable fiber to be fermented by SCFAs-producing bacteria;

iii) butyrate producing bacteria such as *Odoribacter, Butyricimonas* and *Ruminococcus*;

iv) specific food sources included in ketogenic diet that may directly provide the adequate amount of butyric acid such as dairy foods (butter and cheese);

v) fermented foods (kefir, yogurt, tempeh), naturally enriched in SCFAs [168-171].

As a matter of fact, butter is one of the richest butyric acid food sources with an inherent natural supply of 3-4% of fat content as butyric acid. For example, one tablespoon of butter is composed of 560 mg of butyric acid [172]. Thus, for individuals following a ketogenic diet, it is easily possible to consume well more than 1000 mg of butyrate in a day, from natural sources [172]. Hence, butyrate acts in synergy with the ketogenic goals since it represents a direct substrate to undergo beta-oxidation [173]. In line with these concepts, Nagpal and colleagues observed a slight increase in fecal butyrate after 6-weeks of modified Mediterranean-ketogenic diet. The authors supposed that the butyrate might have originated in the liver as consequence of the ketogenic state, or the ketogenic diet might have promoted the intestinal production of butyrate by supplying plant-based fermentable fibers to be fermented by bacteria [169]. Notably, it should be also underlined that our KEMEPHY was

composed also of functional fermented products (kefir, kimchi, whole yogurt and fermented cheese) which are naturally enriched in short-chain fatty acids [170]. In addition, beta-hydroxybutyrate derived from hepatic production during ketogenesis, has also the ability to influence, directly or indirectly, the gut microbiome, providing additional support for the fundamental function of ketone bodies at both intestinal and systemic level [141].

Current limitations

Despite these interesting results, our study is not without limitations. First, the reduced sample size of our cohort of athletes may represent a limit for a real robust statistical difference in gut microbiome profiling. Moreover, our analysis has been performed with 16S rRNA gene sequencing which represent the most applied method to investigating gut microbiome, but it is not efficient as shotgun metagenomic sequencing [174]. Indeed, 16s rRNA targets and reads a region of the 16S rRNA gene while shotgun technique sequences all given genomic DNA while achieving strain-level resolution. The results is that 16S rRNA gene sequencing detects only part of the gut microbiome community revealed by shotgun sequencing and it does not provide a functional profiling of gut microbes [175]. However, a technical challenge was considerable at the time of analysis. Since our research was conducted there years ago and shotgun metagenomic was orders of magnitude more expensive and relatively new than amplicon analysis (~$150 USD for shotgun and ~$50 USD for 16S), at that moment, 16S rRNA sequencing represented the best and most used method for microbiome studies. Moreover, it is important to highlight that also regular physical exercise, such as that performed by our cohort of semi-professional soccer players, might have influenced the results of the study by promoting the maintenance of a functional and physiological microbiota in both groups [176]. Further studies on KD on athletes would help validate these findings in gut microbiome and, thanks to the innovative available bioinformatic platforms, the integration of omics-data with the metagenomic methods may improve the understanding of the relationship between diet, gut microbiome and physical exercise [2]. In addition, our study did not measure the level of SCFAs that could be an

additional finding helping the explanation of the underlying mechanisms and of the interpretation of results.

Conclusion

There is a growing body of research on the role of gut microbiome in sport and performance. For the first time our results demonstrate that i) KEMEPHY diet may be considered a feasible and safe nutritional strategy for athletes to get an adequate body composition, ii) KEMEPHY diet do not change the overall composition of gut microbiome and, iii) thirty days of KEMEPHY intervention may represent an alternative tool for maintaining and/or modulating the composition of gut microbiome in athletes practicing regular exercise. These findings suggest that KEMEPHY diet may represent an efficient dietary pattern for athletes, according to the notion that preserving a stable gut microbiome during dietary intervention represent a marker of gut health and greater athletic performance. It should be stressed that our KEMEPHY diet was mainly composed by healthy fats (good sources of monounsaturated and polyunsaturated fats), fibers (low-carb veggies, seeds), plant-based protein (tofu, tempeh) and fermented foods (kefir, tempeh, yogurt, kimchi), different from a standard high fat - low fibers ketogenic diet, which may not arouse the same beneficial effects on gut microbiome. Our findings demonstrate also that changes in microbial taxa pre and post intervention significantly correlate with environmental variables such as athlete's macronutrient intake. Finally, it should be emphasized that data analysis performed with not updated database may give back partially different results as we demonstrated here.

References:

1. Sommer, F., et al., *The resilience of the intestinal microbiota influences health and disease.* Nat Rev Microbiol, 2017. **15**(10): p. 630-638.

2. Mancin, L., et al., *Optimizing Microbiota Profiles for Athletes.* Exerc Sport Sci Rev, 2021. **49**(1): p. 42-49.

3. Mancin, L., G.D. Wu, and A. Paoli, *Gut microbiota-bile acid-skeletal muscle axis.* Trends Microbiol, 2022.

4. Paoli, A., et al., *Ketogenic Diet and Microbiota: Friends or Enemies?* Genes (Basel), 2019. **10**(7).

5. Sorrenti, V., et al., *Cocoa Polyphenols and Gut Microbiota Interplay: Bioavailability, Prebiotic Effect, and Impact on Human Health.* Nutrients, 2020. **12**(7).

6. Mancin, L., et al., *Effect of 30 days of ketogenic Mediterranean diet with phytoextracts on athletes' gut microbiome composition.* Front Nutr, 2022. **9**: p. 979651.

7. Ley, R.E., et al., *Microbial ecology: human gut microbes associated with obesity.* Nature, 2006. **444**(7122): p. 1022-3.

8. Human Microbiome Project, C., *Structure, function and diversity of the healthy human microbiome.* Nature, 2012. **486**(7402): p. 207-14.

9. Sender, R., S. Fuchs, and R. Milo, *Revised Estimates for the Number of Human and Bacteria Cells in the Body.* PLoS Biol, 2016. **14**(8): p. e1002533.

10. Sanger, F. and A.R. Coulson, *A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase.* J Mol Biol, 1975. **94**(3): p. 441-8.

11. Eckburg, P.B., et al., *Diversity of the human intestinal microbial flora.* Science, 2005. **308**(5728): p. 1635-8.

12. Walker, A.W., et al., *Phylogeny, culturing, and metagenomics of the human gut microbiota.* Trends Microbiol, 2014. **22**(5): p. 267-74.

13. Lane, D.J., et al., *Rapid determination of 16S ribosomal RNA sequences for phylogenetic analyses.* Proc Natl Acad Sci U S A, 1985. **82**(20): p. 6955-9.

14. Fukuda, K., et al., *Molecular Approaches to Studying Microbial Communities: Targeting the 16S Ribosomal RNA Gene.* J UOEH, 2016. **38**(3): p. 223-32.

15. Handelsman, J., *Metagenomics: application of genomics to uncultured microorganisms.* Microbiol Mol Biol Rev, 2004. **68**(4): p. 669-85.

16. Kuczynski, J., et al., *Experimental and analytical tools for studying the human microbiome.* Nat Rev Genet, 2011. **13**(1): p. 47-58.

17. Segata, N., et al., *Computational meta'omics for microbial community studies.* Mol Syst Biol, 2013. **9**: p. 666.

18. Lewandowska, D.W., et al., *Unbiased metagenomic sequencing complements specific routine diagnostic methods and increases chances to detect rare viral strains.* Diagn Microbiol Infect Dis, 2015. **83**(2): p. 133-8.

19. Junemann, S., et al., *Bioinformatics for NGS-based metagenomics and the application to biogas research.* J Biotechnol, 2017. **261**: p. 10-23.

20. Xia, Y. and J. Sun, *Hypothesis Testing and Statistical Analysis of Microbiome.* Genes Dis, 2017. **4**(3): p. 138-148.

21. Bokulich, N.A., et al., *Associations among Wine Grape Microbiome, Metabolome, and Fermentation Behavior Suggest Microbial Contribution to Regional Wine Characteristics.* mBio, 2016. **7**(3).

22. Boisseau, N., N. Barnich, and C. Koechlin-Ramonatxo, *The Nutrition-Microbiota-Physical Activity Triad: An Inspiring New Concept for Health and Sports Performance.* Nutrients, 2022. **14**(5).

23. Wilkinson, D.J., M. Piasecki, and P.J. Atherton, *The age-related loss of skeletal muscle mass and function: Measurement and physiology of muscle fibre atrophy and muscle fibre loss in humans.* Ageing Res Rev, 2018. **47**: p. 123-132.

24. Schirmer, M., et al., *Linking the Human Gut Microbiome to Inflammatory Cytokine Production Capacity.* Cell, 2016. **167**(4): p. 1125-1136 e8.

25. Fu, J., et al., *The Gut Microbiome Contributes to a Substantial Proportion of the Variation in Blood Lipids.* Circ Res, 2015. **117**(9): p. 817-24.

26. Zeevi, D., et al., *Personalized Nutrition by Prediction of Glycemic Responses.* Cell, 2015. **163**(5): p. 1079-1094.

27. McIver, L.J., et al., *bioBakery: a meta'omic analysis environment.* Bioinformatics, 2018. **34**(7): p. 1235-1237.

28. Quince, C., et al., *Corrigendum: Shotgun metagenomics, from sampling to analysis.* Nat Biotechnol, 2017. **35**(12): p. 1211.

29. Langmead, B. and S.L. Salzberg, *Fast gapped-read alignment with Bowtie 2.* Nat Methods, 2012. **9**(4): p. 357-9.

30. Nielsen, H.B., et al., *Identification and assembly of genomes and genetic elements in complex metagenomic samples without using reference genomes.* Nat Biotechnol, 2014. **32**(8): p. 822-8.

31. Mulligan, A.A., et al., *A new tool for converting food frequency questionnaire data into nutrient and food group values: FETA research methods and availability.* BMJ Open, 2014. **4**(3): p. e004503.

32. Bowyer, R.C.E., et al., *Use of dietary indices to control for diet in human gut microbiota studies.* Microbiome, 2018. **6**(1): p. 77.

33. Drescher, L.S., S. Thiele, and G.B. Mensink, *A new index to measure healthy food diversity better reflects a healthy diet than traditional measures.* J Nutr, 2007. **137**(3): p. 647-51.

34. Satija, A., et al., *Healthful and Unhealthful Plant-Based Diets and the Risk of Coronary Heart Disease in U.S. Adults.* J Am Coll Cardiol, 2017. **70**(4): p. 411-422.

35. Fung, T.T., et al., *Diet-quality scores and plasma concentrations of markers of inflammation and endothelial dysfunction.* Am J Clin Nutr, 2005. **82**(1): p. 163-73.

36. Xia, G.H., et al., *Stroke Dysbiosis Index (SDI) in Gut Microbiome Are Associated With Brain Injury and Prognosis of Stroke.* Front Neurol, 2019. **10**: p. 397.

37. Johnson, A.J., et al., *Daily Sampling Reveals Personalized Diet-Microbiome Associations in Humans.* Cell Host Microbe, 2019. **25**(6): p. 789-802 e5.

38. Arumugam, M., et al., *Enterotypes of the human gut microbiome.* Nature, 2011. **473**(7346): p. 174-80.

39. Wu, G.D., et al., *Linking long-term dietary patterns with gut microbial enterotypes.* Science, 2011. **334**(6052): p. 105-8.

40. Falony, G., et al., *Population-level analysis of gut microbiome variation.* Science, 2016. **352**(6285): p. 560-4.

41. Biomarkers Definitions Working, G., *Biomarkers and surrogate endpoints: preferred definitions and conceptual framework.* Clin Pharmacol Ther, 2001. **69**(3): p. 89-95.

42. Schwieterman, M.L., et al., *Strawberry flavor: diverse chemical compositions, a seasonal influence, and effects on sensory perception.* PLoS One, 2014. **9**(2): p. e88446.

43. Rothschild, D., et al., *Environment dominates over host genetics in shaping human gut microbiota.* Nature, 2018. **555**(7695): p. 210-215.

44. Tett, A., et al., *The Prevotella copri Complex Comprises Four Distinct Clades Underrepresented in Westernized Populations.* Cell Host Microbe, 2019. **26**(5): p. 666-679 e7.

45. Petersen, L.M., et al., *Community characteristics of the gut microbiomes of competitive cyclists.* Microbiome, 2017. **5**(1): p. 98.

46. Hooper, L., et al., *Effects of chocolate, cocoa, and flavan-3-ols on cardiovascular health: a systematic review and meta-analysis of randomized trials.* Am J Clin Nutr, 2012. **95**(3): p. 740-51.

47. van Dam, R.M., N. Naidoo, and R. Landberg, *Dietary flavonoids and the development of type 2 diabetes and cardiovascular diseases: review of recent findings.* Curr Opin Lipidol, 2013. **24**(1): p. 25-33.

48. Raman, G., et al., *Dietary intakes of flavan-3-ols and cardiovascular health: a field synopsis using evidence mapping of randomized trials and prospective cohort studies.* Syst Rev, 2018. **7**(1): p. 100.

49. Mao, T., et al., *Cocoa procyanidins and human cytokine transcription and secretion.* J Nutr, 2000. **130**(8S Suppl): p. 2093S-9S.

50. Mao, T.K., et al., *Modulation of TNF-alpha secretion in peripheral blood mononuclear cells by cocoa flavanols and procyanidins.* Dev Immunol, 2002. **9**(3): p. 135-41.

51. Kuebler, U., et al., *Dark chocolate attenuates intracellular pro-inflammatory reactivity to acute psychosocial stress in men: A randomized controlled trial.* Brain Behav Immun, 2016. **57**: p. 200-208.

52. Davinelli, S., et al., *Short-term supplementation with flavanol-rich cocoa improves lipid profile, antioxidant status and positively influences the AA/EPA ratio in healthy subjects.* J Nutr Biochem, 2018. **61**: p. 33-39.

53. Saini, R.K. and Y.S. Keum, *Omega-3 and omega-6 polyunsaturated fatty acids: Dietary sources, metabolism, and significance - A review.* Life Sci, 2018. **203**: p. 255-267.

54. Calder, P.C., *n-3 polyunsaturated fatty acids, inflammation, and inflammatory diseases.* Am J Clin Nutr, 2006. **83**(6 Suppl): p. 1505S-1519S.

55. McGlory, C., P.C. Calder, and E.A. Nunes, *The Influence of Omega-3 Fatty Acids on Skeletal Muscle Protein Turnover in Health, Disuse, and Disease.* Front Nutr, 2019. **6**: p. 144.

56. Lewis, N.A., et al., *Are There Benefits from the Use of Fish Oil Supplements in Athletes? A Systematic Review.* Adv Nutr, 2020. **11**(5): p. 1300-1314.

57. Heaton, L.E., et al., *Selected In-Season Nutritional Strategies to Enhance Recovery for Team Sport Athletes: A Practical Overview.* Sports Med, 2017. **47**(11): p. 2201-2218.

58. Metherel, A.H., et al., *Docosahexaenoic acid is both a product of and a precursor to tetracosahexaenoic acid in the rat.* J Lipid Res, 2019. **60**(2): p. 412-420.

59. Davinelli, S., et al., *Relationship Between Distance Run Per Week, Omega-3 Index, and Arachidonic Acid (AA)/Eicosapentaenoic Acid (EPA) Ratio: An Observational Retrospective Study in Non-elite Runners.* Front Physiol, 2019. **10**: p. 487.

60. Ritz, P.P., et al., *Dietary and Biological Assessment of the Omega-3 Status of Collegiate Athletes: A Cross-Sectional Analysis.* PLoS One, 2020. **15**(4): p. e0228834.

61. Walker, R.E., et al., *Predicting the effects of supplemental EPA and DHA on the omega-3 index.* Am J Clin Nutr, 2019. **110**(4): p. 1034-1040.

62. de Lorgeril, M., et al., *Interactions of wine drinking with omega-3 fatty acids in patients with coronary heart disease: a fish-like effect of moderate wine drinking.* Am Heart J, 2008. **155**(1): p. 175-81.

63. Schramm, D.D., et al., *Chocolate procyanidins decrease the leukotriene-prostacyclin ratio in humans and human aortic endothelial cells.* Am J Clin Nutr, 2001. **73**(1): p. 36-40.

64. Davinelli, S., et al., *Cardioprotection by Cocoa Polyphenols and omega-3 Fatty Acids: A Disease-Prevention Perspective on Aging-Associated Cardiovascular Risk.* J Med Food, 2018. **21**(10): p. 1060-1069.

65. Lee, K.W., et al., *Cocoa has more phenolic phytochemicals and a higher antioxidant capacity than teas and red wine.* J Agric Food Chem, 2003. **51**(25): p. 7292-5.

66. Rechner, A.R., et al., *Colonic metabolism of dietary polyphenols: influence of structure on microbial fermentation products.* Free Radic Biol Med, 2004. **36**(2): p. 212-25.

67. Clarke, S.F., et al., *Exercise and associated dietary extremes impact on gut microbial diversity.* Gut, 2014. **63**(12): p. 1913-20.

68. Faul, F., et al., *G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences.* Behav Res Methods, 2007. **39**(2): p. 175-91.

69. Spadafranca, A., et al., *Effect of dark chocolate on plasma epicatechin levels, DNA resistance to oxidative stress and total antioxidant activity in healthy subjects.* Br J Nutr, 2010. **103**(7): p. 1008-14.

70. Katz, D.L., K. Doughty, and A. Ali, *Cocoa and chocolate in human health and disease.* Antioxid Redox Signal, 2011. **15**(10): p. 2779-811.

71. Wan, Y., et al., *Effects of cocoa powder and dark chocolate on LDL oxidative susceptibility and prostaglandin concentrations in humans.* Am J Clin Nutr, 2001. **74**(5): p. 596-602.

72. Mursu, J., et al., *Dark chocolate consumption increases HDL cholesterol concentration and chocolate fatty acids may inhibit lipid peroxidation in healthy humans.* Free Radic Biol Med, 2004. **37**(9): p. 1351-9.

73. Moser, E.C., et al., *Long-term risk of cardiovascular disease after treatment for aggressive non-Hodgkin lymphoma.* Blood, 2006. **107**(7): p. 2912-9.

74. Ruiz, J.R., et al., *[Sports requiring stressful physical exertion cause abnormalities in plasma lipid profile].* Rev Esp Cardiol, 2004. **57**(6): p. 499-506.

75. Aellen, R., W. Hollmann, and U. Boutellier, *Effects of aerobic and anaerobic training on plasma lipoproteins.* Int J Sports Med, 1993. **14**(7): p. 396-400.

76. Eisenmann, J.C., et al., *Blood lipids of young distance runners: distribution and inter-relationships among training volume, peak oxygen consumption, and body fatness.* Eur J Appl Physiol, 2001. **85**(1-2): p. 104-12.

77. Smith, L.L., *Acute inflammation: the underlying mechanism in delayed onset muscle soreness?* Med Sci Sports Exerc, 1991. **23**(5): p. 542-51.

78. Moldoveanu, A.I., R.J. Shephard, and P.N. Shek, *The cytokine response to physical activity and training.* Sports Med, 2001. **31**(2): p. 115-44.

79. Hardardottir, I., C. Grunfeld, and K.R. Feingold, *Effects of endotoxin and cytokines on lipid metabolism.* Curr Opin Lipidol, 1994. **5**(3): p. 207-15.

80. Nonogaki, K., et al., *Interleukin-6 stimulates hepatic triglyceride secretion in rats.* Endocrinology, 1995. **136**(5): p. 2143-9.

81. Costa, R.J.S., et al., *Assessment of Exercise-Associated Gastrointestinal Perturbations in Research and Practical Settings: Methodological Concerns and Recommendations for Best Practice.* Int J Sport Nutr Exerc Metab, 2022. **32**(5): p. 387-418.

82. Feingold, K.R., et al., *Endotoxin, TNF, and IL-1 decrease cholesterol 7 alpha-hydroxylase mRNA levels and activity.* J Lipid Res, 1996. **37**(2): p. 223-8.

83. Jia, L., et al., *Short-term effect of cocoa product consumption on lipid profile: a meta-analysis of randomized controlled trials.* Am J Clin Nutr, 2010. **92**(1): p. 218-25.

84. Tokede, O.A., J.M. Gaziano, and L. Djousse, *Effects of cocoa products/dark chocolate on serum lipids: a meta-analysis.* Eur J Clin Nutr, 2011. **65**(8): p. 879-86.

85. Nishizaki, Y. and H. Daida, *Optimal Dose of n-3 Polyunsaturated Fatty Acids for Cardiovascular Event Prevention.* Circ Rep, 2020. **2**(4): p. 260-264.

86. Suzuki, K., et al., *Characterization and Modulation of Systemic Inflammatory Response to Exhaustive Exercise in Relation to Oxidative Stress.* Antioxidants (Basel), 2020. **9**(5).

87. Clauss, M., et al., *Interplay Between Exercise and Gut Microbiome in the Context of Human Health and Performance.* Front Nutr, 2021. **8**: p. 637010.

88. Fassarella, M., et al., *Gut microbiome stability and resilience: elucidating the response to perturbations in order to modulate gut health.* Gut, 2021. **70**(3): p. 595-605.

89. Furber, M.J.W., et al., *Gut Microbial Stability is Associated with Greater Endurance Performance in Athletes Undertaking Dietary Periodization.* mSystems, 2022. **7**(3): p. e0012922.

90. Stellingwerff, T., J.P. Morton, and L.M. Burke, *A Framework for Periodized Nutrition for Athletics.* Int J Sport Nutr Exerc Metab, 2019. **29**(2): p. 141-151.

91. Maughan, R.J., et al., *IOC consensus statement: dietary supplements and the high-performance athlete.* Br J Sports Med, 2018. **52**(7): p. 439-455.

92. Lloyd-Price, J., G. Abu-Ali, and C. Huttenhower, *The healthy human microbiome.* Genome Med, 2016. **8**(1): p. 51.

93. Wu, G.D., et al., *Comparative metabolomics in vegans and omnivores reveal constraints on diet-dependent gut microbiota metabolite production.* Gut, 2016. **65**(1): p. 63-72.

94. Honda, K. and D.R. Littman, *The microbiome in infectious disease and inflammation.* Annu Rev Immunol, 2012. **30**: p. 759-95.

95. Dambekodi, P.C. and S.E. Gilliland, *Incorporation of cholesterol into the cellular membrane of Bifidobacterium longum.* J Dairy Sci, 1998. **81**(7): p. 1818-24.

96. Kim, Y.T., et al., *In vivo Trial of Bifidobacterium longum Revealed the Complex Network Correlations Between Gut Microbiota and Health Promotional Effects.* Front Microbiol, 2022. **13**: p. 886934.

97. Shin, J.H., et al., *Consumption of 85% cocoa dark chocolate improves mood in association with gut microbial changes in healthy adults: a randomized controlled trial.* J Nutr Biochem, 2022. **99**: p. 108854.

98. Sokol, H., et al., *Faecalibacterium prausnitzii is an anti-inflammatory commensal bacterium identified by gut microbiota analysis of Crohn disease patients.* Proc Natl Acad Sci U S A, 2008. **105**(43): p. 16731-6.

99. Bressa, C., et al., *Differences in gut microbiota profile between women with active lifestyle and sedentary women.* PLoS One, 2017. **12**(2): p. e0171352.

100. O'Donovan, C.M., et al., *Distinct microbiome composition and metabolome exists across subgroups of elite Irish athletes.* J Sci Med Sport, 2020. **23**(1): p. 63-68.

101. Cerda, B., et al., *Gut Microbiota Modification: Another Piece in the Puzzle of the Benefits of Physical Exercise in Health?* Front Physiol, 2016. **7**: p. 51.

102. Le Roy, T., et al., *Dysosmobacter welbionis is a newly isolated human commensal bacterium preventing diet-induced obesity and metabolic disorders in mice.* Gut, 2022. **71**(3): p. 534-543.

103. Hu, H.J., et al., *Obesity Alters the Microbial Community Profile in Korean Adolescents.* PLoS One, 2015. **10**(7): p. e0134333.

104. Tims, S., et al., *Microbiota conservation and BMI signatures in adult monozygotic twins.* ISME J, 2013. **7**(4): p. 707-17.

105. Egshatyan, L., et al., *Gut microbiota and diet in patients with different glucose tolerance.* Endocr Connect, 2016. **5**(1): p. 1-9.

106. Tuovinen, E., et al., *Cytokine response of human mononuclear cells induced by intestinal Clostridium species.* Anaerobe, 2013. **19**: p. 70-6.

107. Pache, I., G. Rogler, and C. Felley, *TNF-alpha blockers in inflammatory bowel diseases: practical consensus recommendations and a user's guide.* Swiss Med Wkly, 2009. **139**(19-20): p. 278-87.

108. Gupta, V.K., S. Paul, and C. Dutta, *Geography, Ethnicity or Subsistence-Specific Variations in Human Microbiome Composition and Diversity.* Front Microbiol, 2017. **8**: p. 1162.

109. Jin, M., et al., *The role of intestinal microbiota in cardiovascular disease.* J Cell Mol Med, 2019. **23**(4): p. 2343-2350.

110. Lynch, S.V. and O. Pedersen, *The Human Intestinal Microbiome in Health and Disease.* N Engl J Med, 2016. **375**(24): p. 2369-2379.

111. David, L.A., et al., *Diet rapidly and reproducibly alters the human gut microbiome.* Nature, 2014. **505**(7484): p. 559-63.

112. Scheiman, J., et al., *Meta-omics analysis of elite athletes identifies a performance-enhancing microbe that functions via lactate metabolism.* Nat Med, 2019. **25**(7): p. 1104-1109.

113. Barton, W., et al., *The microbiome of professional athletes differs from that of more sedentary subjects in composition and particularly at the functional metabolic level.* Gut, 2018. **67**(4): p. 625-633.

114. Poole, A.C., et al., *Human Salivary Amylase Gene Copy Number Impacts Oral and Gut Microbiomes.* Cell Host Microbe, 2019. **25**(4): p. 553-564 e7.

115. Andeweg, S.P., C. Kesmir, and B.E. Dutilh, *Quantifying the Impact of Human Leukocyte Antigen on the Human Gut Microbiota.* mSphere, 2021. **6**(4): p. e0047621.

116. Burke, L.M., et al., *Carbohydrates for training and competition.* J Sports Sci, 2011. **29 Suppl 1**: p. S17-27.

117. Sonnenburg, E.D. and J.L. Sonnenburg, *Starving our microbial self: the deleterious consequences of a diet deficient in microbiota-accessible carbohydrates.* Cell Metab, 2014. **20**(5): p. 779-786.

118. Paoli, A., A. Bianco, and K.A. Grimaldi, *The Ketogenic Diet and Sport: A Possible Marriage?* Exerc Sport Sci Rev, 2015. **43**(3): p. 153-62.

119. Cooder, H.R., *Epilepsy in Children: With Particular Reference to the Ketogenic Diet.* Cal West Med, 1933. **39**(3): p. 169-73.

120. Perez-Guisado, J., *[Ketogenic diets: additional benefits to the weight loss and unfounded secondary effects].* Arch Latinoam Nutr, 2008. **58**(4): p. 323-9.

121. Antonio Paoli, A., et al., *Effects of 30 days of ketogenic diet on body composition, muscle strength, muscle area, metabolism, and performance in semi-professional soccer players.* J Int Soc Sports Nutr, 2021. **18**(1): p. 62.

122. Wolters, M., et al., *Dietary fat, the gut microbiota, and metabolic health - A systematic review conducted within the MyNewGut project.* Clin Nutr, 2019. **38**(6): p. 2504-2520.

123. Murtaza, N., et al., *The Effects of Dietary Pattern during Intensified Training on Stool Microbiota of Elite Race Walkers.* Nutrients, 2019. **11**(2).

124. Kim, D., et al., *Optimizing methods and dodging pitfalls in microbiome research.* Microbiome, 2017. **5**(1): p. 52.

125. Leone, A., et al., *Food and Food Products on the Italian Market for Ketogenic Dietary Treatment of Neurological Diseases.* Nutrients, 2019. **11**(5).

126. Paoli, A., et al., *Long term successful weight loss with a combination biphasic ketogenic Mediterranean diet and Mediterranean diet maintenance protocol.* Nutrients, 2013. **5**(12): p. 5205-17.

127. Paoli, A., et al., *Effects of a ketogenic diet in overweight women with polycystic ovary syndrome.* J Transl Med, 2020. **18**(1): p. 104.

128. CJ, D.C.H., et al., *The Effect of Medium Chain Triglycerides on Time to Nutritional Ketosis and Symptoms of Keto-Induction in Healthy Adults: A Randomised Controlled Clinical Trial.* J Nutr Metab, 2018. **2018**: p. 2630565.

129. Jager, R., et al., *International Society of Sports Nutrition Position Stand: protein and exercise.* J Int Soc Sports Nutr, 2017. **14**: p. 20.

130. Black, K.E., P.M. Skidmore, and R.C. Brown, *Energy intakes of ultraendurance cyclists during competition, an observational study.* Int J Sport Nutr Exerc Metab, 2012. **22**(1): p. 19-23.

131. Bolyen, E., et al., *Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2.* Nat Biotechnol, 2019. **37**(8): p. 852-857.

132. Callahan, B.J., et al., *DADA2: High-resolution sample inference from Illumina amplicon data.* Nat Methods, 2016. **13**(7): p. 581-3.

133. Quast, C., et al., *The SILVA ribosomal RNA gene database project: improved data processing and web-based tools.* Nucleic Acids Res, 2013. **41**(Database issue): p. D590-6.

134. Ding, R.X., et al., *Revisit gut microbiota and its impact on human health and disease.* J Food Drug Anal, 2019. **27**(3): p. 623-631.

135. Zmora, N., J. Suez, and E. Elinav, *You are what you eat: diet, health and the gut microbiota.* Nat Rev Gastroenterol Hepatol, 2019. **16**(1): p. 35-56.

136. Singh, R.P., et al., *High-Fat Diet Induced Alteration of Mice Microbiota and the Functional Ability to Utilize Fructooligosaccharide for Ethanol Production.* Front Cell Infect Microbiol, 2020. **10**: p. 376.

137. Daniel, H., et al., *High-fat diet alters gut microbiota physiology in mice.* ISME J, 2014. **8**(2): p. 295-308.

138. Xiao, L., et al., *High-fat feeding rather than obesity drives taxonomical and functional changes in the gut microbiota in mice.* Microbiome, 2017. **5**(1): p. 43.

139. Serino, M., et al., *Metabolic adaptation to a high-fat diet is associated with a change in the gut microbiota.* Gut, 2012. **61**(4): p. 543-53.

140. Morrison, K.E., et al., *It's the fiber, not the fat: significant effects of dietary challenge on the gut microbiome.* Microbiome, 2020. **8**(1): p. 15.

141. Ang, Q.Y., et al., *Ketogenic Diets Alter the Gut Microbiome Resulting in Decreased Intestinal Th17 Cells.* Cell, 2020. **181**(6): p. 1263-1275 e16.

142. Caesar, R., et al., *Crosstalk between Gut Microbiota and Dietary Lipids Aggravates WAT Inflammation through TLR Signaling.* Cell Metab, 2015. **22**(4): p. 658-68.

143. Huang, E.Y., et al., *Composition of dietary fat source shapes gut microbiota architecture and alters host inflammatory mediators in mouse adipose tissue.* JPEN J Parenter Enteral Nutr, 2013. **37**(6): p. 746-54.

144. Costantini, L., et al., *Impact of Omega-3 Fatty Acids on the Gut Microbiota.* Int J Mol Sci, 2017. **18**(12).

145. Paoli, A., et al., *Effects of n-3 polyunsaturated fatty acids (omega-3) supplementation on some cardiovascular risk factors with a ketogenic Mediterranean diet.* Mar Drugs, 2015. **13**(2): p. 996-1009.

146. Pokusaeva, K., G.F. Fitzgerald, and D. van Sinderen, *Carbohydrate metabolism in Bifidobacteria.* Genes Nutr, 2011. **6**(3): p. 285-306.

147. Turroni, F., et al., *Glycan Utilization and Cross-Feeding Activities by Bifidobacteria.* Trends Microbiol, 2018. **26**(4): p. 339-350.

148. Grill, J.P., J. Crociani, and J. Ballongue, *Characterization of fructose 6 phosphate phosphoketolases purified from Bifidobacterium species.* Curr Microbiol, 1995. **31**(1): p. 49-54.

149. Ellulu, M.S., et al., *Obesity and inflammation: the linking mechanism and the complications.* Arch Med Sci, 2017. **13**(4): p. 851-863.

150. Sainsbury, E., et al., *Effect of dietary carbohydrate restriction on glycemic control in adults with diabetes: A systematic review and meta-analysis.* Diabetes Res Clin Pract, 2018. **139**: p. 239-252.

151. Volek, J.S., et al., *Carbohydrate restriction has a more favorable impact on the metabolic syndrome than a low fat diet.* Lipids, 2009. **44**(4): p. 297-309.

152. Santacruz, A., et al., *Interplay between weight loss and gut microbiota composition in overweight adolescents.* Obesity (Silver Spring), 2009. **17**(10): p. 1906-15.

153. De Palma, G., et al., *Effects of a gluten-free diet on gut microbiota and immune function in healthy adult human subjects.* Br J Nutr, 2009. **102**(8): p. 1154-60.

154. Lin, D., et al., *Association of dietary fibre intake and gut microbiota in adults.* Br J Nutr, 2018. **120**(9): p. 1014-1022.

155. Magne, F., et al., *The Firmicutes/Bacteroidetes Ratio: A Relevant Marker of Gut Dysbiosis in Obese Patients?* Nutrients, 2020. **12**(5).

156. de Wit, N., et al., *Saturated fat stimulates obesity and hepatic steatosis and affects gut microbiota composition by an enhanced overflow of dietary fat to the distal intestine.* Am J Physiol Gastrointest Liver Physiol, 2012. **303**(5): p. G589-99.

157. De Filippo, C., et al., *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, 2010. **107**(33): p. 14691-6.

158. Hildebrandt, M.A., et al., *High-fat diet determines the composition of the murine gut microbiome independently of obesity.* Gastroenterology, 2009. **137**(5): p. 1716-24 e1-2.

159. Verdam, F.J., et al., *Human intestinal microbiota composition is associated with local and systemic inflammation in obesity.* Obesity (Silver Spring), 2013. **21**(12): p. E607-15.

160. Alexopoulos, N., D. Katritsis, and P. Raggi, *Visceral adipose tissue as a source of inflammation and promoter of atherosclerosis.* Atherosclerosis, 2014. **233**(1): p. 104-12.

161. Goodrich, J.K., et al., *Human genetics shape the gut microbiome.* Cell, 2014. **159**(4): p. 789-99.

162. McNulty, N.P., et al., *Effects of diet on resource utilization by a model human gut microbiota containing Bacteroides cellulosilyticus WH2, a symbiont with an extensive glycobiome.* PLoS Biol, 2013. **11**(8): p. e1001637.

163. Garcia-Mantrana, I., et al., *Shifts on Gut Microbiota Associated to Mediterranean Diet Adherence and Specific Dietary Intakes on General Adult Population.* Front Microbiol, 2018. **9**: p. 890.

164. Donati Zeppa, S., et al., *Nine weeks of high-intensity indoor cycling training induced changes in the microbiota composition in non-athlete healthy male college students.* J Int Soc Sports Nutr, 2021. **18**(1): p. 74.

165. Koh, A., et al., *From Dietary Fiber to Host Physiology: Short-Chain Fatty Acids as Key Bacterial Metabolites.* Cell, 2016. **165**(6): p. 1332-1345.

166. Lai, Z.L., et al., *Fecal microbiota transplantation confers beneficial metabolic effects of diet and exercise on diet-induced obese mice.* Sci Rep, 2018. **8**(1): p. 15625.

167. Lim, M.Y., et al., *The effect of heritability and host genetics on the gut microbiota and metabolic syndrome.* Gut, 2017. **66**(6): p. 1031-1038.

168. Zhai, S., et al., *Dietary butyrate suppresses inflammation through modulating gut microbiota in high-fat diet-fed mice.* FEMS Microbiol Lett, 2019. **366**(13).

169. Nagpal, R., et al., *Modified Mediterranean-ketogenic diet modulates gut microbiome and short-chain fatty acids in association with Alzheimer's disease markers in subjects with mild cognitive impairment.* EBioMedicine, 2019. **47**: p. 529-542.

170. Wolfe, B.E. and R.J. Dutton, *Fermented foods as experimentally tractable microbial ecosystems.* Cell, 2015. **161**(1): p. 49-55.

171. Shimizu, H., et al., *Dietary short-chain fatty acid intake improves the hepatic metabolic condition via FFAR3.* Sci Rep, 2019. **9**(1): p. 16574.

172. Cavaleri, F. and E. Bashar, *Potential Synergies of beta-Hydroxybutyrate and Butyrate on the Modulation of Metabolism, Inflammation, Cognition, and General Health.* J Nutr Metab, 2018. **2018**: p. 7195760.

173. Hird, F.J. and R.H. Symons, *The mechanism of ketone-body formation from butyrate in rat liver.* Biochem J, 1962. **84**: p. 212-6.

174. Durazzi, F., et al., *Comparison between 16S rRNA and shotgun sequencing data for the taxonomic characterization of the gut microbiota.* Sci Rep, 2021. **11**(1): p. 3030.

175. Quince, C., et al., *Shotgun metagenomics, from sampling to analysis.* Nat Biotechnol, 2017. **35**(9): p. 833-844.

176. Monda, V., et al., *Exercise Modifies the Gut Microbiota with Positive Health Effects.* Oxid Med Cell Longev, 2017. **2017**: p. 3831972.

Publications:

Mancin L.,* Gary D. Wu., Paoli A., Gut microbiota–bile acid–skeletal muscle axis, Trends in Microbiology, DOI: 10.1016/j.tim.2022.10.003., 2022 **IF: 18.23**

Mancin L.,* Amatori S., Caprio M., Sattin E., Bertoldi L., Paoli A., Effect of 30 days of ketogenic Mediterranean diet with phytoextracts on athletes' gut microbiome composition, Front in Nutr., DOI:10.3389/fnut.2022.979651., 2022 **IF: 6.576**

de Paiva AKF, de Oliveira EP, **Mancin L**, Paoli A, Mota JF. Effects of probiotic supplementation on performance of resistance and aerobic exercises: a systematic review. Nutr Rev. 2022 Aug 11:nuac046. doi: 10.1093/nutrit/nuac046. Epub ahead of print. PMID: 35950956. IF:**7.11**

Mancin L., * Rollo I., Mota J.F., Piccini F., Carletti M., Susto G., Valle G., Paoli A. Optimizing Microbiota Profiles for Athletes: Dream or Reality? Exerc Sport Sci Rev. 2020 oct 9. DOI: 10.1249/jes.0000000000000236. Epub ahead of print. Pmid: 33044333. **IF: 4.370**

Antonio Paoli, A., **Mancin, L**., Caprio, M. et al. Effects of 30 days of ketogenic diet on body composition, muscle strength, muscle area, metabolism, and performance in semi-professional soccer players. J Int Soc Sports Nutr 18, 62 (2021). https://doi.org/10.1186/s12970-021-00459-9 **IF:4.832**

Paoli A., **Mancin L**., Bianco A., Thomas E., Mota J.F., Piccini F., Ketogenic Diet and Microbiota: Friends or Enemies? Genes (Basel). 2019 jul 15;10(7):534. DOI: 10.3390/genes10070534. Pmid: 31311141; pmcid: pmc6678592. IF: **2.988**

Paoli A., **Mancin L.,** Giacona M.C., Bianco A., Caprio M., Effects of Ketogenic Diet in Overweight Women with Polycystic Ovary Dyndrome. J Transl Med. 2020 feb 27;18(1):104. DOI: 10.1186/s12967-020-02277-0. Pmid: 32103756; pmcid: pmc7045520. **IF: 4.197**

Ashtary-Larky D, Bagheri R, Bavi H, Baker JS, Moro T, **Mancin L**, Paoli A. Ketogenic diets, physical activity, and body composition: A review. Br J Nutr. 2021 Jul 12:1-68. doi: 10.1017/S0007114521002609. Epub ahead of print. PMID: 34250885. **IF: 3.52**

Paoli A., **Mancin L**., Saoncella M., Grigoletto D., Pacelli F.Q., Zamparo P., Schoenfeld B.J., Marcolin G., Mind-muscle Connection: Effects of Verbal Instructions on Muscle Activity During Bench Press Exercise. Eur j transl myol. 2019 jun 12;29(2):8250. DOI: 10.4081/ejtm.2019.8250. Pmid: 31354928; pmcid: pmc6615069. **IF: 3.781**

Sorrenti V., Ali S., **Mancin L**., Davinelli S., Paoli A., Scapagnini G., Cocoa Polyphenols and Gut Microbiota Interplay: Bioavailability, Prebiotic effect, and Impact on Human Health. Nutrients. 2020 jun 27;12(7):1908. DOI: 10.3390/nu12071908. Pmid: 32605083; pmcid: pmc7400387. **IF: 4.546**