

## Article

# Gut Microbiota and Serum Metabolome in Elite Cross-Country Skiers: A Controlled Study

Jukka E. Hintikka <sup>1,\*</sup>, Eveliina Munukka <sup>2</sup>, Maarit Valtonen <sup>3</sup>, Raakel Luoto <sup>4</sup>, Johanna K. Ihalainen <sup>1</sup>, Teemu Kallonen <sup>5</sup>, Matti Waris <sup>6</sup>, Olli J. Heinonen <sup>7</sup>, Olli Ruuskanen <sup>4</sup> and Satu Pekkala <sup>1</sup>

<sup>1</sup> Faculty of Sport and Health Sciences, University of Jyväskylä, 40014 Jyväskylä, Finland; johanna.k.ihalainen@jyu.fi (J.K.I.); satpekka@jyu.fi (S.P.)

<sup>2</sup> Turku Microbiome Biobank, Institute of Biomedicine, University of Turku, 20500 Turku, Finland; laevmu@utu.fi

<sup>3</sup> Research Institute for Olympic Sports, 40700 Jyväskylä, Finland; maarit.valtonen@kihu.fi

<sup>4</sup> Department of Pediatrics and Adolescent Medicine, Turku University Hospital, 20521 Turku, Finland; rajopi@utu.fi (R.L.); olli.ruuskanen@tyks.fi (O.R.)

<sup>5</sup> Clinical Microbiology, Turku University Hospital, 20521 Turku, Finland; tevika@utu.fi

<sup>6</sup> Institute of Biomedicine, University of Turku, 20500 Turku, Finland; mwaris@utu.fi

<sup>7</sup> Paavo Nurmi Centre, Department of Health and Physical Activity, University of Turku, 20540 Turku, Finland; olli.heinonen@utu.fi

\* Correspondence: juenhint@jyu.fi

**Abstract:** Exercise has been shown to affect gut the microbiome and metabolic health, with athletes typically displaying a higher microbial diversity. However, research on the gut microbiota and systemic metabolism in elite athletes remains scarce. In this study, we compared the gut microbiota profiles and serum metabolome of national team cross-country skiers at the end of an exhausting training and competitive season to those of normally physically-active controls. The gut microbiota were analyzed using 16S rRNA amplicon sequencing. Serum metabolites were analyzed using nuclear magnetic resonance. Phylogenetic diversity and the abundance of several mucin-degrading gut microbial taxa, including *Akkermansia*, were lower in the athletes. The athletes had a healthier serum lipid profile than the controls, which was only partly explained by body mass index. *Butyrivibrio* associated positively with HDL cholesterol, HDL2 cholesterol and HDL particle size. The *Ruminococcus torques* group was less abundant in the athlete group and positively associated with total cholesterol and VLDL and LDL particles. We found the healthier lipid profile of elite athletes to co-occur with known health-beneficial gut microbes. Further studies should elucidate these links and whether athletes are prone to mucin depletion related microbial changes during the competitive season.

**Keywords:** winter games; athletes; exercise; microbiology; metabolomics; lipids



**Citation:** Hintikka, J.E.; Munukka, E.; Valtonen, M.; Luoto, R.; Ihalainen, J.K.; Kallonen, T.; Waris, M.; Heinonen, O.J.; Ruuskanen, O.; Pekkala, S. Gut Microbiota and Serum Metabolome in Elite Cross-Country Skiers: A Controlled Study. *Metabolites* **2022**, *12*, 335. <https://doi.org/10.3390/metabo12040335>

Academic Editor: Phillipe Soucaille

Received: 4 March 2022

Accepted: 4 April 2022

Published: 7 April 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Gut microbiota refer to the trillions of microbial cells inhabiting the gastrointestinal tract that, in addition to many other tasks, break down the macromolecules and nutrients from ingested food [1,2]. Bacterial metabolites such as short chain fatty acids (SCFAs) provide energy for muscles and intestinal epithelial cells [3–5]. Physical activity and lean body composition are associated with a gut microbiome that contains high abundances of health-promoting bacterial taxa [6–9], and higher microbial diversity is a common finding in athletes compared to sedentary controls [10]. Exercise may help develop a microbiome with a greater ability to harness energy from the diet and with an increased capacity for carbohydrate metabolism, cell structure, and nucleotide biosynthesis [11,12]. On the other hand, dysbiotic, i.e., metabolically unbalanced, gut microbiota have been shown to impair skeletal muscle adaptation to exercise [13]. Certain gut microbes have been shown to increase and diminish acutely in response to exercise [14], and high-intensity endurance training has even been linked to increased dysbiosis on some occasions [15].

During training and competition season, elite-level athletes are predisposed to increased levels of energy expenditure, stress and low-grade inflammation [16,17] and are at an increased risk of respiratory infections [18]. Chronic training can have notable effects on metabolic and inflammatory states [19], and these effects can persist in workloads below those constituting overtraining [20,21]. To this end, new studies suggest that certain microbial metabolic functions can enhance athletic performance and these functions could be accessed with certain bacterial taxa: Scheiman et al. [22] demonstrated that inoculation into mice of the lactate-fermenting species *Veillonella atypica*, isolated from Boston marathon competitors, increased the run times of the mice on a treadmill. Likewise, a recent similar trial on humans demonstrated that oral administration of a subspecies of *Bifidobacterium longum* isolated from a weightlifting Olympic athlete could increase Cooper's running test results [23].

Both physical activity and the gut microbiome can have an instrumental role in maintaining cardiovascular and metabolic health [24–27], yet chronic high-intensity exercise can also increase inflammation and cause shifts in the microbiome. Integrative studies on the gut microbiome and the interplay of microbial and systemic metabolism in elite athletes are still scarce. We investigated whether the gut microbiota and serum metabolome of high-level cross-country skiers at the end of the competition season differs from those of age and sex-matched non-athletes.

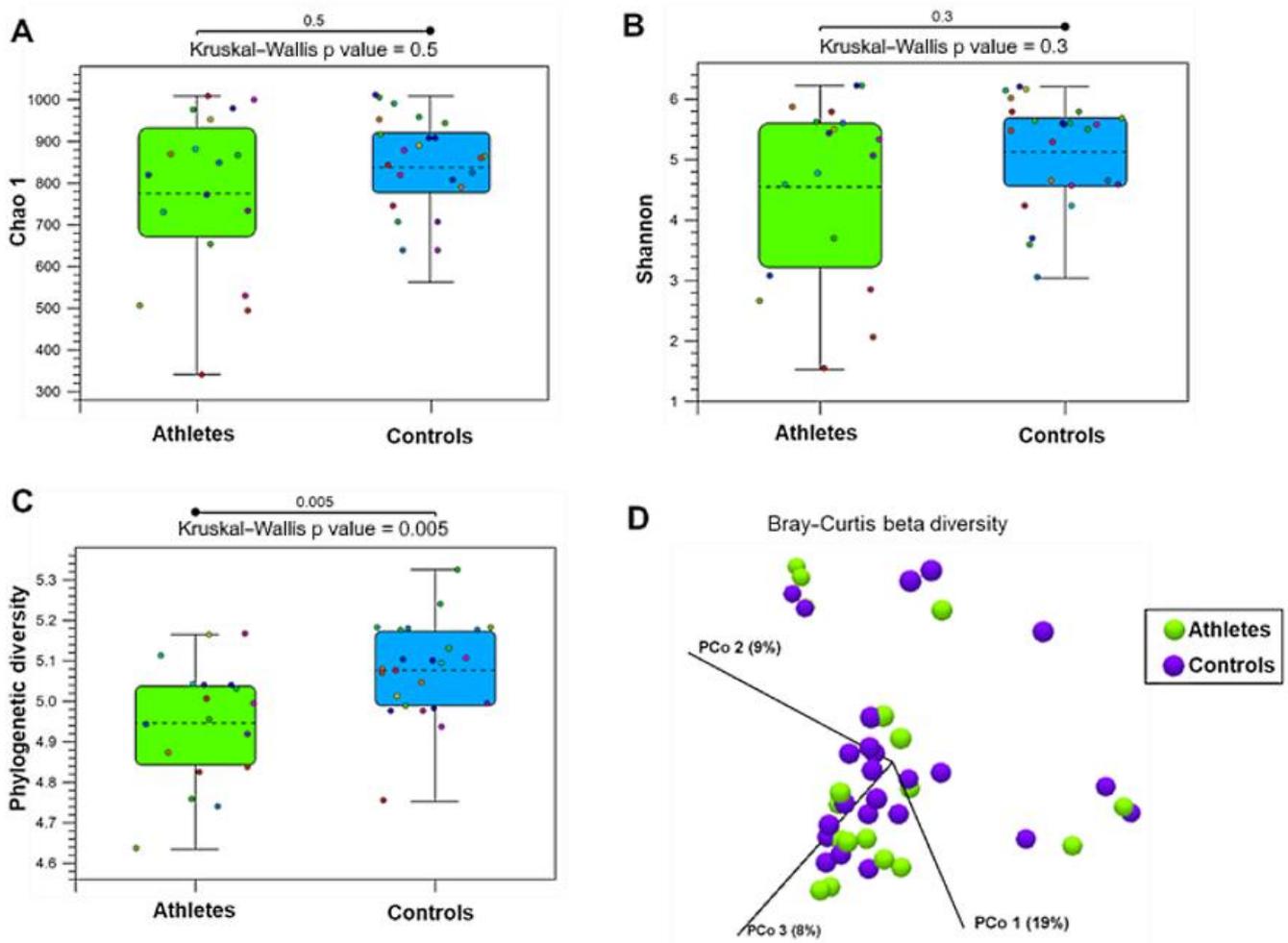
## 2. Results

### 2.1. Gut Microbiota Diversity

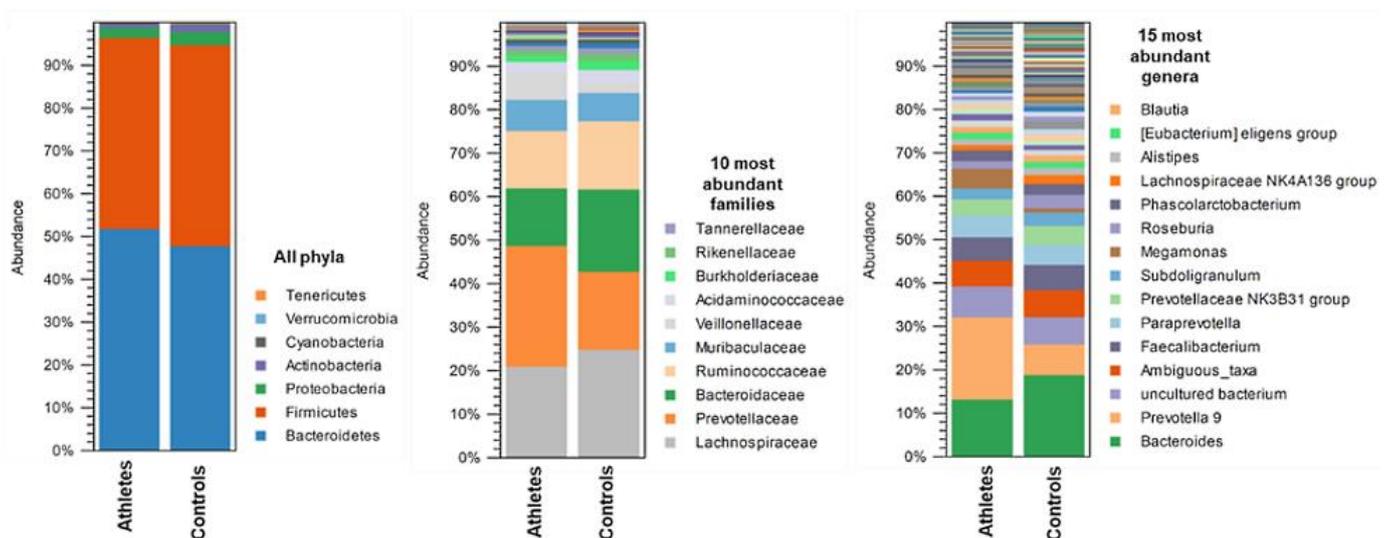
The gut microbiota alpha-diversity measures, Chao1 (i.e., species richness, Figure 1A) and Shannon index (i.e., species diversity, Figure 1B) were found to be similar among the athletes and controls. However, the athletes had lower phylogenetic diversity than the controls (Figure 1C). According to the Bray–Curtis distance and PERMANOVA analysis ( $p = 0.66$ ), the groups did not differ in beta-diversity, that is, in inter-individual species diversity of the gut microbiota (Figure 1D).

### 2.2. Gut Microbiota Composition

The average gut microbiota composition of the athletes and controls at the phylum, family and genus levels are shown in Figure 2. In the athletes, Bacteroidetes (50.4% of all sequences) and Firmicutes (46.0%) were the dominant bacterial phyla, followed by Proteobacteria (2.3%), Actinobacteria (0.79%), Verrucomicrobia (0.05%), Cyanobacteria (0.20%) and Tenericutes (0.03%). In the controls, Firmicutes (48.3%) and Bacteroidetes (46.2% of all sequences) were the dominant bacterial phyla, followed by Proteobacteria (3.36%), Actinobacteria (1.57%), Cyanobacteria (0.33%), Verrucomicrobia (0.20%), and Tenericutes (0.14%). A total of 27 families and 82 genera were identified in the athletes and controls. The family *Bacteroidae* (13.3%, mostly genus *Bacteroides*) explained the dominance of Bacteroidetes in the athletes, and the family *Lachnospiraceae* (25.5%, mostly genus *Blautia*) the dominance of Firmicutes in the controls.

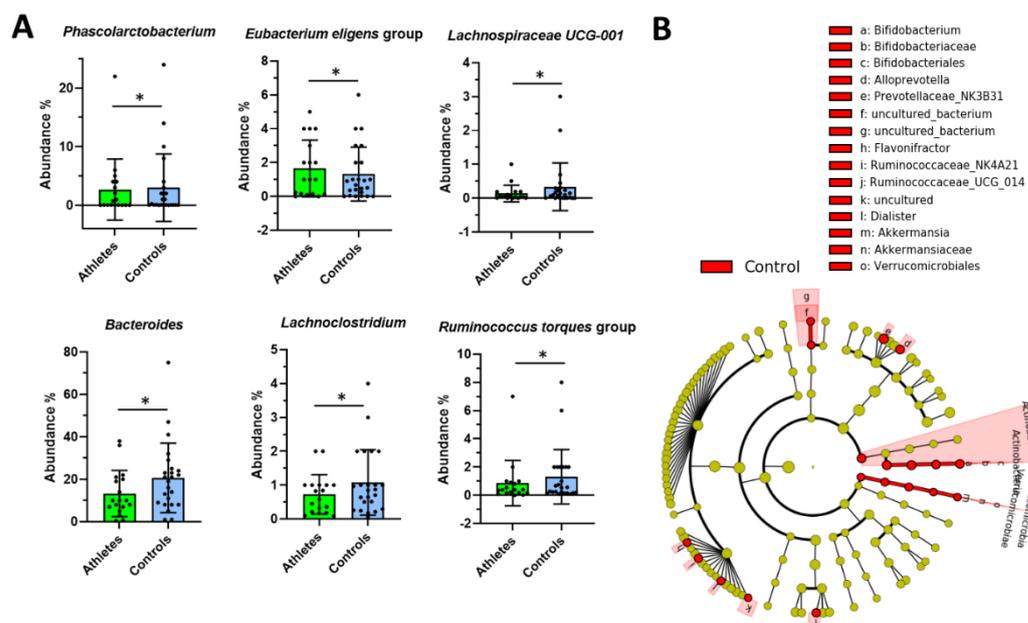


**Figure 1.** The diversity of the gut microbiota in the athletes ( $n = 27$ ) and controls ( $n = 27$ ). (A) Chao1, a measure of alpha-diversity indicating species richness. (B) Shannon index, a measure of alpha-diversity indicating species diversity. (C) Phylogenetic diversity. (D) Principal coordination (PCo) plot of beta-diversity analyzed by Bray-Curtis distance.



**Figure 2.** The average gut microbiota composition in the athletes ( $n = 27$ ) and controls ( $n = 27$ ) at phylum, family, and genus level.

In the ANOVA-like comparison, differences in the abundance of six gut microbiota genera were found between the groups (Figure 3). The athletes had a lower relative abundance of *Phascolarctobacterium*, *Lachnospiraceae* UCG-001, *Bacteroides*, *Lachnoclostridium* and the *Ruminococcus torques* group as well as a higher relative abundance of the *Eubacterium eligens* group than the controls ( $p < 0.038$  for all). However, LEfSe analysis, that takes into account biological consistency and effect size, revealed the phylum *Actinobacteria* and the genera *Akkermansia*, *Bifidobacterium*, the *Prevotellaceae* NK3B31 group, *Alloprevotella*, *Flavonifractor*, *Ruminococcaceae* UCG 014 and the *Ruminococcaceae* NK4A214 group, to be more abundant in the controls.

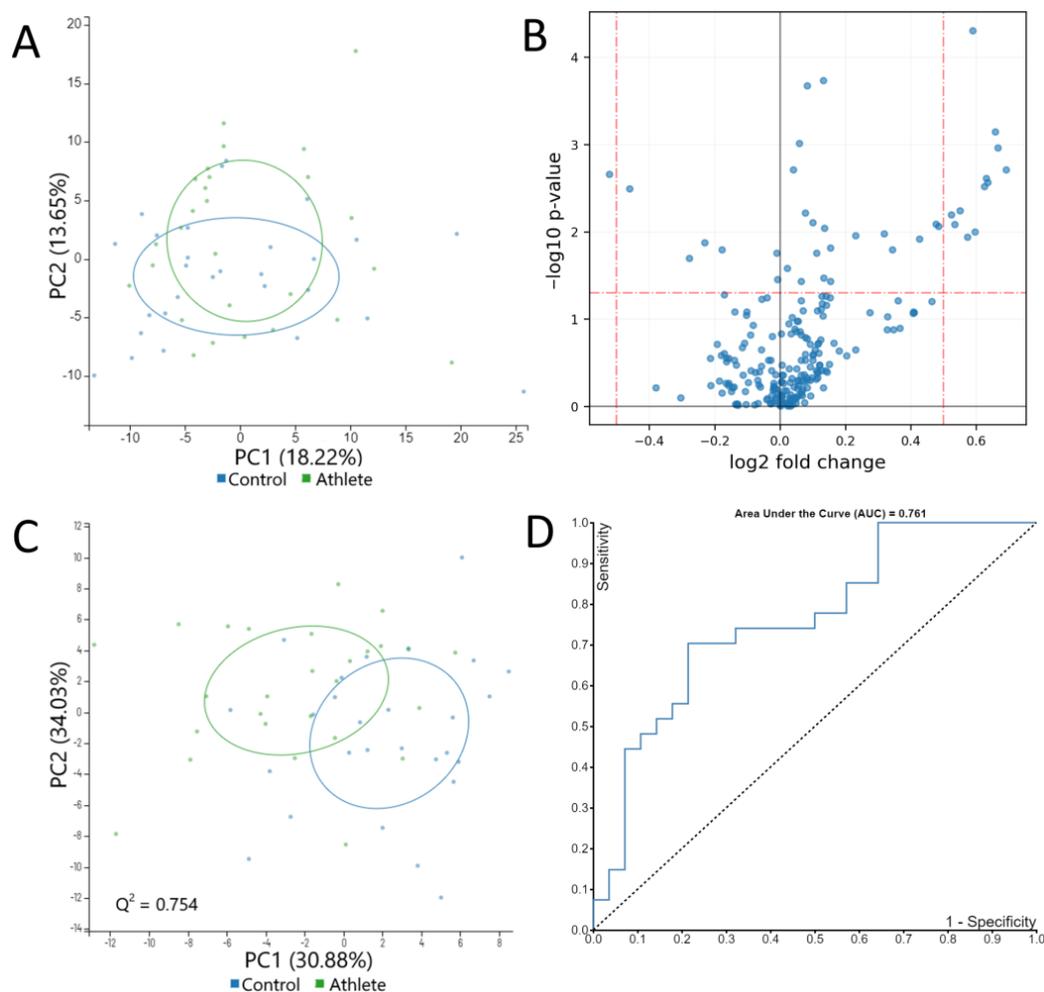


**Figure 3.** (A) The differential genus level abundances of the gut microbiota in the athletes ( $n = 18$ ) and controls ( $n = 24$ ). ANOVA-like comparison  $p$  values  $* < 0.05$ . (B) Cladogram from the LEfSe-analysis. The gut microbiota taxa that are highlighted were found to be enriched in the control group.

In the control group, the amount of weekly exercise was associated positively with the family *Pasteurellaceae* and inversely with *Acidamidococcaceae* (Supplements, Figure S1). The training load of the athletes was associated inversely with *Enterobacteriaceae*, *Bacteroidaceae* and *Veillonellaceae*. In addition, age was inversely correlated with *Tannerellaceae* and body mass index (BMI) with *Clostridium* sp. K4410.MGS-306.

### 2.3. Serum Metabolome

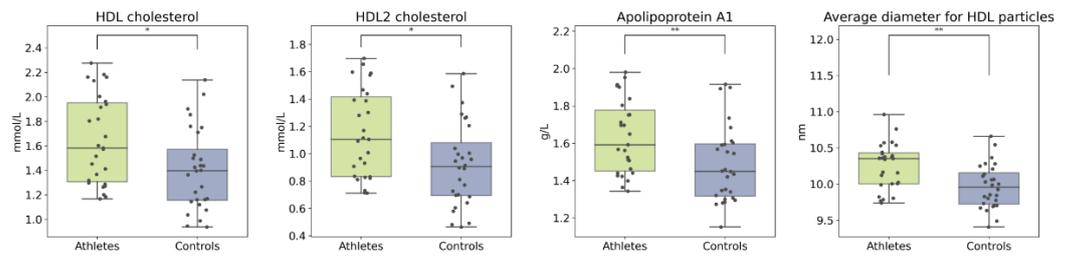
Overall, the athletes and controls were largely similar in their serum metabolome. Over the two first principal components, principal component analysis (PCA) showed no separation between the groups (Figure 4A). The supervised method, partial least squares-discriminant analysis (PLS-DA), reached a predictive ability of 0.75 (Q squared) and showed a moderate level of separation between the groups (Figure 4B). None of the metabolites differed by large factors (Figure 4C). A random forest classification task reached a sensitivity of 0.7 and a specificity of approximately 0.8 at peak (Figure 4D).



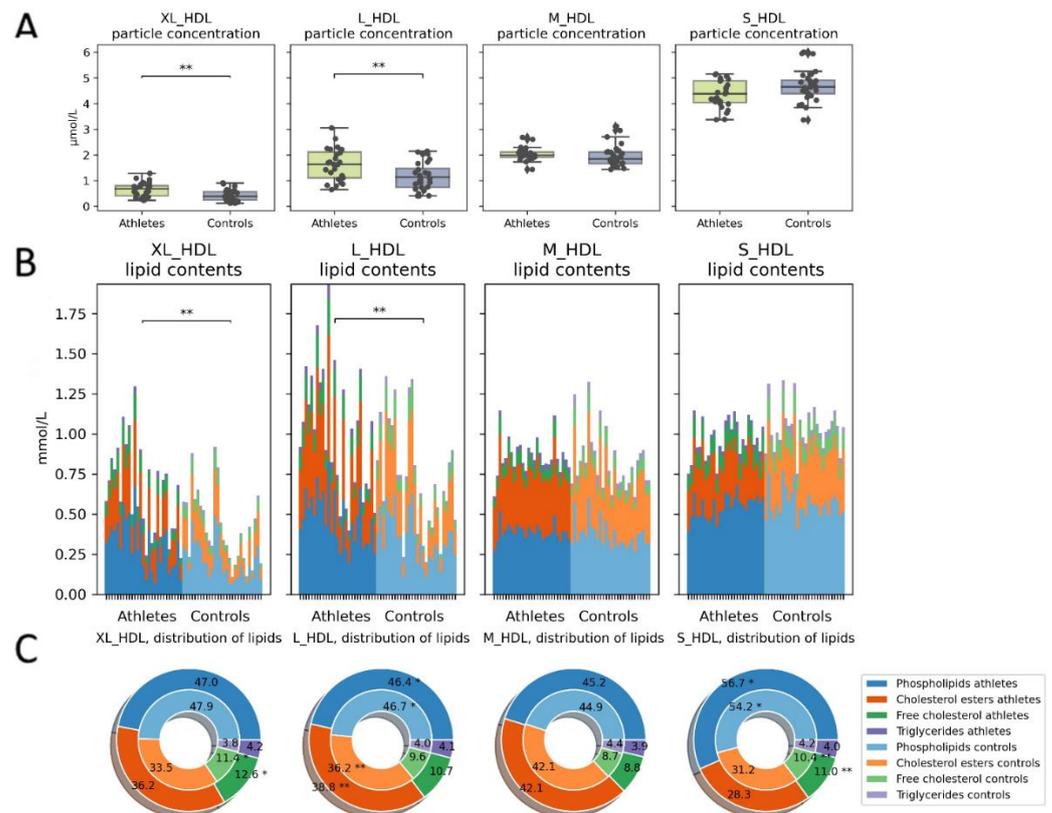
**Figure 4.** (A) The principal component (PC) analysis plot of all metabolites shows that most of the variance was explained by other factors than group. (B) The volcano plot indicates that the group differences were less than 2-fold for all metabolites. (C) The partial least squares-discriminant analysis plot and (D) receiver operand characteristics curve for random forest show that the classification task reached moderate accuracy.

The age of the participants associated positively with circulating lipids including total cholesterol, total apolipoproteins, and LDL particle concentrations. BMI associated negatively with HDL cholesterol, average HDL size, amino acids, and glycolysis-related metabolites (Supplements, Figure S2). Despite similarities between the groups, the athletes had higher total concentrations of both total HDL and its subfraction HDL2 compared to controls (Figure 5). Apolipoprotein A, the backbone of HDL, was higher and the mean HDL particle size was larger in the athlete group. In addition, the ratio of saturated fatty acids to total fatty acids was significantly elevated in the athletes. After adjusting for age and BMI, the group differences remained significant (see Figure S3).

The absolute and relative amounts of lipids within the different-sized HDL particles varied both within and between the groups. The athletes had a higher concentration of large and very large HDL particles (Figure 6A) and, consequently, higher lipid concentrations in these particles (Figure 6B). In the large and very large HDL particles (typically corresponding to HDL2 subfractions), the athletes had a higher cholesterol to phospholipid ratio than the controls did (Figure 6C). This difference was reversed as the particle size decreased.



**Figure 5.** Boxplots of HDL cholesterol and related metabolites. Kruskal–Wallis  $p$  value: \* < 0.05, \*\* < 0.01.



**Figure 6.** HDL particle concentrations and contents. (A) Total particle concentrations and (B) absolute lipid concentrations in the particles. Athletes had more large to very large HDL particles and consequently more lipids contained within them. (C) Relative proportions of lipids: the outer donut chart represents the athletes, and the inner represents controls. The athletes had a smaller phospholipid/cholesterol ratio in the large particles than the controls did. Kruskal–Wallis  $p$  value: \* < 0.05, \*\* < 0.01.

The primary ketone bodies beta-hydroxybutyrate (bOHB) and acetoacetate were slightly lower in the athlete group, as was glycerol. In addition, the athletes had a higher concentration of pyruvate (Figure 7). After adjusting for age and BMI, the group differences in these metabolites remained significant (see Supplements: Figure S3).





BMI, branched chain amino acids and four microbial taxa (*Blautia*, *Dorea*, *Ruminococcus*, and SHA-98).

In addition, a set of studies has explored the gut microbiomes of professional rugby players compared to healthy controls [11,32]. The microbiome was both compositionally and functionally different between the groups and strongly associated with various aspects of diet. In contrast to our results, the alpha- and phylogenetic diversities of the gut microbiota were found to be higher and *Akkermansia*, a SCFA-producing and health-beneficial genus of the phylum *Verrucomicrobia*, more abundant in the athletes. The findings of the rugby study could partly be explained by dietary extremes observed in the groups, whereas we did not analyze diet. In another study comparing senior orienteers to sedentary controls, no differences in the microbial diversity were found altogether [33]. Previously, microbial diversity and the abundance of *Akkermansia* have been positively linked with physical activity in several other studies [6]. Since *Akkermansia* did not associate with weekly physical activity in either of our study groups, our results somewhat contradict these findings. However, it should be noted that our control group was not sedentary but exercised normally.

The genus *R. torques* group, which was more abundant in the controls than it was in the athletes (1.3% vs. 0.8%, respectively), has been associated with increased intestinal barrier leakage and increased serum triglycerides [34,35]. Although *R. torques* group did not associate with serum triglycerides in our study, we found the genus, along with *Lachnospiraceae* UCG-008, to be positively associated with total cholesterol, small VLDLs and large LDLs, which are mediators in hyperlipidemia [36,37]. These genera had an inverse association with serum acetate, which is both a microbial metabolite and a less ubiquitous ketone body. Interestingly, the species *R. torques*, along with *Akkermansia muciniphila* and species of *Bacteroides* and *Bifidobacterium*, is a consumer of mucins, the glycosylated proteins in the intestinal mucosa [38]. These genera were also less abundant in the athletes, which could indicate lower mucin availability, possibly due to high physical stress during the competitive season. However, this hypothesis warrants further studies.

*Butyricoccus* is a butyrate-producing genus of the gut bacteria that has been proposed as a next generation probiotic [39]. The most prominent species of this genus, *B. pullicaecorum* has been shown to inversely associate with the incidences of ulcerative colitis and Crohn's disease [40], suggesting that it can be health-beneficial. Ketone bodies acetoacetate and bOHB can be used as a substrate by several butyrate-producing bacteria [41,42] including *Butyricoccus*. Here, the relative abundance of *Butyricoccus* was not higher in the athlete group, but acetoacetate inversely and significantly predicted its abundance, which might indicate utilization of this ketone body by *Butyricoccus*. Further, this genus was positively associated with higher HDL and HDL2 cholesterol and larger HDL particle size. This association has not been observed before but is an encouraging discovery considering future studies and raises the question whether the species belonging to this genus could mediate a healthier blood lipid profile when orally administered as probiotics or, vice versa, whether they thrive under such conditions.

The HDL cholesterol associates with improved health outcomes [43,44], and analyzing lipoprotein particle size and different HDL subfractions allows for more accurate probing of lipoproteins and associated health risks [45–47]. Mature HDL particles that are formed from small, lipid-poor, pre-beta HDLs are further divided into subfractions HDL2 and HDL3 according to their density and size [45]. Of the subfractions, HDL2 are larger, less dense and more lipid-rich forms of HDL. Both subfractions transport cholesterol into the liver to be excreted as bile salts, but HDL2 is also able to excrete cholesterol directly via transintestinal cholesterol excretion [44]. In our study, we found the athletes to have a higher proportion of cholesterol in the largest HDLs and, conversely, a lower proportion in the smallest ones. This could indicate more efficient cholesterol excretion in the athletes either as bile salts or through transintestinal cholesterol excretion. The gut microbes can both deconjugate bile salts and assimilate cholesterol [48], and some bacteria, including *R.*

*Torques*, readily convert primary bile acids into secondary [49]. This could in part explain the observed associations between bacterial genera and lipoproteins.

A limitation of this study is the difficulty in selecting a control group matched for body composition. No body composition measurements were performed in the study and BMI alone can give a limited insight into this. Dietary data was not recorded in this study, and although not the primary interest in the study, the diet can affect both the gut microbiota and lipid parameters. Generally, elite athletes can be presumed to eat a different diet from the general population, e.g., a diet rich in carbohydrates and protein [50], while an individual diet can vary depending on many factors [51]. The absence of overnight fasting can confound the concentrations of some metabolic markers such as total triglycerides, yet this heterogeneity affected both groups equally. Since our study is cross-sectional, all results are associative and are interpreted as such. When comparing athletes and the general population, genetic background can confound some of the associations since the same polygenic risk factors can predict both physical activity and health outcomes [52,53]. Likewise, the gut microbiome exhibits at least some degree of heritability [54]. The possible ergogenic capacity (e.g., ability to ferment lactate) in the gut microbiome of elite skiers should be elucidated with further studies.

To conclude, we found lower phylogenetic diversity in elite cross-country skiers as well as minor overall differences in the gut microbiota composition and serum metabolome between the athletes and controls. The athletes had lower abundances of several bacterial genera, including many mucin-degrading bacteria, and a healthier serum lipid profile. *Butyricoccus*, a genus with potential as a next-generation probiotic, was associated with higher HDL cholesterol and larger HDL particle size. In addition, we found the *R. torques* group, a genus associated with gastrointestinal disorders and dyslipidemia, less abundant in the athlete group, and associated with more abundant VLDL and LDL particles. Whether the athletes' serum lipid profile is facilitated by the microbiome or vice versa, or whether these factors share a genetic makeup, could be studied with metagenomics and fecal metabolomics. Further longitudinal studies could also investigate whether elite athletes during competitive season are prone to mucin depletion and related changes in microbial composition and diversity.

## 4. Materials and Methods

### 4.1. Study Design and Population

This observational case-control study was carried out during the Finnish Nordic Ski Championships in Äänekoski, Finland between 28 March and 1 April 2019, an event held at the end of the skiers' competition year. The study recruitment was accepted by 27 of 28 athletes belonging to the national Nordic Ski Team of Finland. The training season of the athletes had started at the beginning of May and the competition season at the beginning of November in 2018. The athletes had, therefore, experienced heavy physical stress for 11 months. For every athlete, one healthy, moderately exercising (<6 h per week) control subject was recruited from among the students and staff of Turku University Hospital and University of Turku, Finland. The controls ( $n = 27$ ) were matched for age ( $\pm 2$  years) and sex. The control subjects were studied in Turku, Finland according to the same study protocol as the athletes in Äänekoski, between 2 and 11 April 2019.

The clinical data and health-related information from the athletes and controls were collected by interview at the study visit by the study nurse (Table 1). The training load of the athletes was collected from the day-to-day training diary data of the previous 11 months.

The blood samples were collected by the study nurse a day before the competition. The athletes had followed their own individual training protocol when preparing for the following day's race. The time frame between the last training bout and the blood sampling varied and we did not record the exact time. Due to the competition circumstances, no fasting from athletes and, correspondingly, nor from controls was required. The blood samples were centrifuged, and the serum separated immediately. The serum was aliquoted, immediately frozen at  $-20$  °C and then stored at  $-80$  °C.

**Table 1.** Clinical characteristics of the study population. Values are represented as mean  $\pm$  SD or as *n* of participants.

	Athletes ( <i>n</i> = 27)	Controls ( <i>n</i> = 27)
Age	27.1 $\pm$ 5.1	27.4 $\pm$ 5.6
Male/Female	14/13 (52/48%)	14/13 (52/48%)
BMI	22.05 $\pm$ 1.8	24.0 $\pm$ 3.5
Exercise load, h/week	15 $\pm$ 2	4 $\pm$ 1
Used oral antibiotics during previous 6 weeks	2	1

BMI = Body mass index.

#### 4.2. Fecal Sample Collection and Microbial DNA Extraction

Both athletes and control subjects received a package containing self-collection equipment and instructions for fecal sampling. The package also included a questionnaire focusing on health status and lifestyle factors at the time of fecal sampling. The participants were guided to send both the sample and questionnaire to the Microbiome Biobank laboratory (University of Turku, Finland) by mail as soon as possible after the collection. The specimens were collected into OMNIgene<sup>®</sup>•GUT collection tubes (DNA Genotek, Kanata, ON, Canada) according to the manufacturer's instructions. Briefly, the participants were guided to collect a small amount (approximately 500 mg) of fecal material into the tube, to homogenize the sample by vigorous shaking for 30 s and to mark the date and time of the sampling on the accompanying collection form. As OMNIgene<sup>®</sup>•GUT collection tubes include a stabilizing solution that guarantees DNA integrity in typical ambient temperature fluctuations and stability at room temperature for as long as 60 days, collection, storage and shipping of the samples could be performed at ambient temperatures [55].

At the laboratory, the samples were homogenized by gentle mixing, and the bacterial DNA was extracted from 200–250  $\mu$ L of sample solution with a GXT Stool Extraction Kit VER 2.0 (Hain Lifescience GmbH, Nehren, Germany). Before the extraction, an additional homogenization by bead-beating in 1.4 mm Ceramic Bead Tubes (MO BIO Laboratories, Inc., Carlsbad, CA, USA) at 1000 rpm for 3 minutes with a MO BIO PowerLyzer<sup>™</sup> 24 Bench Top Bead-Based Homogenizer (MO BIO Laboratories, Inc., Carlsbad, CA, USA) to enhance the cell lysis. The DNA concentrations were measured with a Qubit dsDNA HS Assay kit and Qubit 2.0 fluorometer (Thermo Fisher Scientific, Waltham, MA, USA), and the DNA samples were stored at  $-75$  °C.

#### 4.3. 16S rRNA Gene Sequencing and Sequence Data Processing

The gut microbiota profiles were analyzed by 16S rRNA gene sequencing. To this end, the variable region V4 of the bacterial 16S rRNA gene was amplified with custom-designed dual-indexed primers and sequenced with the Illumina MiSeq system as described [56]. The raw 16S rRNA gene sequencing data were demultiplexed and the sequence adapters, primers and barcodes were clipped by using the Illumina BaseSpace platform. The raw sequence quality was checked with FastQC [57].

#### 4.4. Gut Microbiota Composition Analyses

The 16S rRNA gene sequences were clustered to operational taxonomic units (OTUs) at 97% similarity using CLC Microbial Genomics Package (Qiagen, Hilden, Germany). The rRNA gene sequences were classified using the SILVA SSU Reference database (v132, 99%). The statistical analyses were performed with CLC Microbial Genomics Package. To analyze the gut microbiota alpha-diversity, Chao1 and Shannon indices were quantified. In addition, phylogenetic diversity was determined. The differences in the alpha-diversity measures between the groups were analyzed with the Kruskal–Wallis test. The beta-diversity analysis, which describes the dissimilarities in the ecosystem level community composition between samples, was based on Bray–Curtis distance and PERMANOVA for significance testing and the group differences were visualized with Principal Coordination Analysis (PCoA)

in CLC. The taxonomic differences between the groups were analyzed with ANOVA-like comparison in CLC Microbial Genomics Package. The statistical significance in group comparisons was set at  $p < 0.05$  after the multiple testing correction (Benjamini–Hochberg false discovery rate). The differences were also analyzed by LEfSe using the browser module and standard protocol [58]. The statistical significance was set at  $p < 0.05$ .

#### 4.5. Metabolite Analyses

A high-throughput proton NMR metabolomics platform (Nightingale Health Ltd., Helsinki, Finland) was used to analyze the serum metabolic profiles as described earlier [59,60]. The analysis platform assesses 228 variables, including biomarkers of lipid and glucose metabolism, amino acids and ketone bodies. Zero values were imputed as the lowest values in the data set above zero, as per the laboratory instructions. Since more than half of the metabolites and derived parameters were not normally distributed, the Kruskal–Wallis test was used to test for feature-wise group differences. The effects of the background variables were investigated using Quade’s ANCOVA (nonparametric analysis of covariance). PCA and PLS-DA were used for multivariate analysis and visualization of the metabolites. To reduce the dimensionality in select visualizations, the metabolites were clustered and summed according to the biological function as reported by Nightingale Health.

For the correlation analyses between the microbes and metabolites, the microbial taxa with a prevalence of less than 10% at 0.001 relative abundance were excluded and the data was center log ratio (clr) transformed. Before the clr transformation, a pseudo-count of  $10^{-8}$  was added to all fields to mitigate zero values. Any metabolites with missing values were also filtered out. The Spearman correlations between the metabolites and microbial taxa were measured and clustered on the microbial axis using the Ward method and Euclidean distance. The general linear model with ordinary least squares method was used to further test the associations. In addition, random forest and *XGBoost* in Python were used to test the predictive ability of the metabolites and taxa on the group membership.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/metabo12040335/s1>: Figure S1. Spearman correlation coefficients between the bacterial families and background variables; Figure S2. Spearman correlation coefficients between the metabolite groups and background variables; Figure S3. The metabolites and derived parameters. Table S1. Results summary for linear regression model using *Butyricoccus* and *Ruminococcus* torques group as dependent variables and associated metabolites as regressors.

**Author Contributions:** M.V., J.K.I. and O.J.H. conceived and designed the analysis. R.L., O.R. and M.W. contributed to the design of the study. E.M. and T.K. designed the collection and analysis of gut microbiota data. J.E.H. and S.P. performed the analysis. J.E.H. wrote the primary manuscript. O.R. and S.P. acquired funding for the study. All authors contributed to the editing of the manuscript, have read, and approved the final version of the manuscript, and agree with the order of presentation of the authors. All authors have read and agreed to the published version of the manuscript.

**Funding:** J.E.H. and S.P. were financially supported by the Academy of Finland Researcher fellowship for S.P. (Grant ID 308042). The analyses of this manuscript were made with funding from the Jenny and Antti Wihuri Foundation.

**Institutional Review Board Statement:** The study complied with the Declaration of Helsinki as revised in 2000, and all study-related activities were conducted according to Good Clinical Practice. The study protocol received a positive statement from the Ethics Committee of the Hospital District of Southwest Finland (ETMK Dnro: 5/1801/2019).

**Informed Consent Statement:** Written informed consent was obtained from all study subjects prior to measurements or sample collections.

**Data Availability Statement:** The data presented in this study are available in article and supplementary material.

**Acknowledgments:** We thank Nightingale Health Ltd. (Helsinki) for performing the NMR metabolomics analysis. We also thank Minna Hyppönen, Wilma Grönroos and Sirpa Vääntinen for technical assistance

with study subjects. Finally, we thank Minna Lamppu for the technical assistance in the gut microbiota sample collection and handling.

**Conflicts of Interest:** E.M. is currently working as full-time Medical Advisor for Biocodex Nordics. The other authors declare no conflict of interest.

## References

1. Qin, J.; Li, R.; Raes, J.; Arumugam, M.; Burgdorf, K.S.; Manichanh, C.; Nielsen, T.; Pons, N.; Levenez, F.; Yamada, T.; et al. A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* **2010**, *464*, 59–65. [\[CrossRef\]](#)
2. Sender, R.; Fuchs, S.; Milo, R. Revised Estimates for the Number of Human and Bacteria Cells in the Body. *PLoS Biol.* **2016**, *14*, e1002533. [\[CrossRef\]](#)
3. Salvi, P.S.; Cowles, R.A. Butyrate and the Intestinal Epithelium: Modulation of Proliferation and Inflammation in Homeostasis and Disease. *Cells* **2021**, *10*, 1775. [\[CrossRef\]](#) [\[PubMed\]](#)
4. Chambers, E.S.; Byrne, C.S.; Aspey, K.; Chen, Y.; Khan, S.; Morrison, D.J.; Frost, G. Acute oral sodium propionate supplementation raises resting energy expenditure and lipid oxidation in fasted humans. *Diabetes Obes. Metab.* **2018**, *20*, 1034–1039. [\[CrossRef\]](#) [\[PubMed\]](#)
5. Bongiovanni, T.; Yin, M.O.L.; Heaney, L. The Athlete and Gut Microbiome: Short-chain Fatty Acids as Potential Ergogenic Aids for Exercise and Training. *Int. J. Sports Med.* **2021**, *42*, 1143–1158. [\[CrossRef\]](#) [\[PubMed\]](#)
6. Lensu, S.; Pekkala, S. Gut Microbiota, Microbial Metabolites and Human Physical Performance. *Metabolites* **2021**, *11*, 716. [\[CrossRef\]](#)
7. Cox, A.J.; West, N.P.; Cripps, A.W. Obesity, inflammation, and the gut microbiota. *Lancet Diabetes Endocrinol.* **2015**, *3*, 207–215. [\[CrossRef\]](#)
8. Mohr, A.E.; Jäger, R.; Carpenter, K.C.; Kerksick, C.M.; Purpura, M.; Townsend, J.R.; West, N.P.; Black, K.; Gleeson, M.; Pyne, D.B.; et al. The athletic gut microbiota. *J. Int. Soc. Sports Nutr.* **2020**, *17*, 1–33. [\[CrossRef\]](#)
9. Dorelli, B.; Gallè, F.; de Vito, C.; Duranti, G.; Iachini, M.; Zaccarin, M.; Standoli, J.P.; Ceci, R.; Romano, F.; Liguori, G.; et al. Can Physical Activity Influence Human Gut Microbiota Composition Independently of Diet? A Systematic Review. *Nutrients* **2021**, *13*, 1890. [\[CrossRef\]](#)
10. Makin, S. Do microbes affect athletic performance? *Nature* **2021**, *592*, S17–S19. [\[CrossRef\]](#)
11. Barton, W.; Penney, N.C.; Cronin, O.; Garcia-Perez, I.; Molloy, M.G.; Holmes, E.; Shanahan, F.; Cotter, P.D.; O’Sullivan, O. The microbiome of professional athletes differs from that of more sedentary subjects in composition and particularly at the functional metabolic level. *Gut* **2018**, *67*, 625–633. [\[CrossRef\]](#)
12. Flint, H.J.; Scott, K.P.; Duncan, S.H.; Louis, P.; Forano, E. Microbial degradation of complex carbohydrates in the gut. *Gut Microbes* **2012**, *3*, 289–306. [\[CrossRef\]](#)
13. Valentino, T.R.; Vechetti, I.J.; Mobley, C.B.; Dungan, C.M.; Golden, L.; Goh, J.; McCarthy, J.J. Dysbiosis of the gut microbiome impairs mouse skeletal muscle adaptation to exercise. *J. Physiol.* **2021**, *599*, 4845–4863. [\[CrossRef\]](#)
14. Zhao, X.; Zhang, Z.; Hu, B.; Huang, W.; Yuan, C.; Zou, L. Response of Gut Microbiota to Metabolite Changes Induced by Endurance Exercise. *Front. Microbiol.* **2018**, *9*, 765. [\[CrossRef\]](#)
15. Morishima, S.; Aoi, W.; Kawamura, A.; Kawase, T.; Takagi, T.; Naito, Y.; Tsukahara, T.; Inoue, R. Intensive, prolonged exercise seemingly causes gut dysbiosis in female endurance runners. *J. Clin. Biochem. Nutr.* **2021**, *68*, 253–258. [\[CrossRef\]](#)
16. Heydenreich, J.; Kayser, B.; Schutz, Y.; Melzer, K. Total Energy Expenditure, Energy Intake, and Body Composition in Endurance Athletes Across the Training Season: A Systematic Review. *Sport. Med.-Open* **2017**, *3*, 1–24. [\[CrossRef\]](#)
17. Koutedakis, Y. Seasonal Variation in Fitness Parameters in Competitive Athletes. *Sport. Med.* **2012**, *19*, 373–392. [\[CrossRef\]](#)
18. Ruuskanen, O.; Luoto, R.; Valtonen, M.; Heinonen, J.O.; Waris, M. Respiratory Viral Infections in Athletes: Many Unanswered Questions. *Sport. Med.* **2022**, 1–9. [\[CrossRef\]](#)
19. Jürimäe, J.; Mäestu, J.; Jürimäe, T.; Mangus, B.; Von Duvillard, S.P. Peripheral signals of energy homeostasis as possible markers of training stress in athletes: A review. *Metabolism* **2011**, *60*, 335–350. [\[CrossRef\]](#)
20. Meeusen, R.; Duclos, M.; Foster, C.; Fry, A.; Gleeson, M.; Nieman, D.; Raglin, J.; Rietjens, G.; Steinacker, J.; Urhausen, A. Prevention, diagnosis and treatment of the overtraining syndrome: Joint consensus statement of the European College of Sport Science (ECSS) and the American College of Sports Medicine (ACSM). *Eur. J. Sport Sci.* **2012**, *13*, 1–24. [\[CrossRef\]](#)
21. Schweltnus, M.; Soligard, T.; Alonso, J.; Bahr, R.; Clarsen, B.; Dijkstra, H.P.; Gabbett, T.; Gleeson, M.; Häggglund, M.; Hutchinson, M.R.; et al. How much is too much? (Part 1) International Olympic Committee consensus statement on load in sport and risk of injury. *Br. J. Sports Med.* **2016**, *50*, 1030–1041. [\[CrossRef\]](#)
22. Scheiman, J.; Lubert, J.M.; Chavkin, T.A.; MacDonald, T.; Tung, A.; Pham, L.D.; Wibowo, M.C.; Wurth, R.C.; Punthambaker, S.; Tierney, B.T.; et al. Meta-omics analysis of elite athletes identifies a performance-enhancing microbe that functions via lactate metabolism. *Nat. Med.* **2019**, *25*, 1104–1109. [\[CrossRef\]](#)
23. Lin, C.L.; Hsu, Y.J.; Ho, H.H.; Chang, Y.C.; Kuo, Y.W.; Yeh, Y.T.; Tsai, S.Y.; Chen, C.W.; Chen, J.F.; Huang, C.C.; et al. Bifidobacterium longum subsp. longum OLP-01 Supplementation during Endurance Running Training Improves Exercise Performance in Middle- and Long-Distance Runners: A Double-Blind Controlled Trial. *Nutrients* **2020**, *12*, 1972. [\[CrossRef\]](#)

24. Wilmot, E.G.; Edwardson, C.L.; Achana, F.A.; Davies, M.J.; Gorely, T.; Gray, L.J.; Khunti, K.; Yates, T.; Biddle, S.J.H. Sedentary time in adults and the association with diabetes, cardiovascular disease and death: Systematic review and meta-analysis. *Diabetologia* **2012**, *55*, 2895–2905. [[CrossRef](#)]
25. Tang, W.H.W.; Hazen, S.L. Microbiome, trimethylamine N-oxide, and cardiometabolic disease. *Transl. Res.* **2017**, *179*, 108–115. [[CrossRef](#)]
26. Utzschneider, K.M.; Kratz, M.; Damman, C.J.; Hullarg, M. Mechanisms Linking the Gut Microbiome and Glucose Metabolism. *J. Clin. Endocrinol. Metab.* **2016**, *101*, 1445–1454. [[CrossRef](#)]
27. Turnbaugh, P.J.; Ley, R.E.; Mahowald, M.A.; Magrini, V.; Mardis, E.R.; Gordon, J.I. An obesity-associated gut microbiome with increased capacity for energy harvest. *Nature* **2006**, *444*, 1027–1031. [[CrossRef](#)]
28. Vojinovic, D.; Radjabzadeh, D.; Kurilshikov, A.; Amin, N.; Wijmenga, C.; Franke, L.; Ikram, M.A.; Uitterlinden, A.G.; Zhernakova, A.; Fu, J.; et al. Relationship between gut microbiota and circulating metabolites in population-based cohorts. *Nat. Commun.* **2019**, *10*, 1–7. [[CrossRef](#)] [[PubMed](#)]
29. Goodrich, J.K.; Waters, J.L.; Poole, A.C.; Sutter, J.L.; Koren, O.; Blehman, R.; Beaumont, M.; Van Treuren, W.; Knight, R.; Bell, J.T.; et al. Human genetics shape the gut microbiome. *Cell* **2014**, *159*, 789–799. [[CrossRef](#)]
30. Dhakal, S.; McCormack, L.; Dey, M. Association of the Gut Microbiota with Weight-Loss Response within a Retail Weight-Management Program. *Microorganisms* **2020**, *8*, 1246. [[CrossRef](#)]
31. Ottosson, F.; Brunkwall, L.; Ericson, U.; Nilsson, P.M.; Almgren, P.; Fernandez, C.; Melander, O.; Orho-Melander, M. Connection Between BMI-Related Plasma Metabolite Profile and Gut Microbiota. *J. Clin. Endocrinol. Metab.* **2018**, *103*, 1491–1501. [[CrossRef](#)] [[PubMed](#)]
32. Clarke, S.F.; Murphy, E.F.; O’Sullivan, O.; Lucey, A.J.; Humphreys, M.; Hogan, A.; Hayes, P.; O’Reilly, M.; Jeffery, I.B.; Wood-Martin, R.; et al. Exercise and associated dietary extremes impact on gut microbial diversity. *Gut* **2014**, *63*, 1913–1920. [[CrossRef](#)] [[PubMed](#)]
33. Fart, F.; Rajan, S.K.; Wall, R.; Rangel, I.; Ganda-Mall, J.P.; Tingö, L.; Brummer, R.J.; Repsilber, D.; Schoultz, I.; Lindqvist, C.M. Differences in Gut Microbiome Composition between Senior Orienteering Athletes and Community-Dwelling Older Adults. *Nutrients* **2020**, *12*, 2610. [[CrossRef](#)] [[PubMed](#)]
34. Rajilić-Stojanović, M.; de Vos, W.M. The first 1000 cultured species of the human gastrointestinal microbiota. *FEMS Microbiol. Rev.* **2014**, *38*, 996–1047. [[CrossRef](#)]
35. Cani, P.D. The gut microbiota manages host metabolism. *Nat. Rev. Endocrinol.* **2013**, *10*, 74–76. [[CrossRef](#)]
36. Horton, J.D.; Shimano, H.; Hamilton, R.L.; Brown, M.S.; Goldstein, J.L. Disruption of LDL receptor gene in transgenic SREBP-1a mice unmasks hyperlipidemia resulting from production of lipid-rich VLDL. *J. Clin. Investig.* **1999**, *103*, 1067–1076. [[CrossRef](#)]
37. Nielsen, S.; Karpe, F. Determinants of VLDL-triglycerides production. *Curr. Opin. Lipidol.* **2012**, *23*, 327. [[CrossRef](#)]
38. Van Herreweghen, F.; De Paepe, K.; Roume, H.; Kerckhof, F.M.; Van de Wiele, T. Mucin degradation niche as a driver of microbiome composition and Akkermansia muciniphila abundance in a dynamic gut model is donor independent. *FEMS Microbiol. Ecol.* **2018**, *94*, 186. [[CrossRef](#)]
39. Boesmans, L.; Valles-Colomer, M.; Wang, J.; Eeckhaut, V.; Falony, G.; Ducatelle, R.; Van Immerseel, F.; Raes, J.; Verbeke, K. Butyrate Producers as Potential Next-Generation Probiotics: Safety Assessment of the Administration of Butyricoccus pullicaecorum to Healthy Volunteers. *mSystems* **2018**, *3*. [[CrossRef](#)]
40. Eeckhaut, V.; Machiels, K.; Perrier, C.; Romero, C.; Maes, S.; Flahou, B.; Steppe, M.; Haesebrouck, F.; Sas, B.; Ducatelle, R.; et al. Butyricococcus pullicaecorum in inflammatory bowel disease. *Gut* **2013**, *62*, 1745–1752. [[CrossRef](#)]
41. Sasaki, K.; Sasaki, D.; Hannya, A.; Tsubota, J.; Kondo, A. In vitro human colonic microbiota utilises D-β-hydroxybutyrate to increase butyrogenesis. *Sci. Rep.* **2020**, *10*, 1–8. [[CrossRef](#)]
42. Paoli, A.; Mancin, L.; Bianco, A.; Thomas, E.; Mota, J.F.; Piccini, F. Ketogenic Diet and Microbiota: Friends or Enemies? *Genes* **2019**, *10*, 534. [[CrossRef](#)]
43. Heinecke, J. HDL and Cardiovascular-Disease Risk—Time for a New Approach? *N. Engl. J. Med.* **2011**, *364*, 170–171. [[CrossRef](#)]
44. Ouimet, M.; Barrett, T.J.; Fisher, E.A. HDL and Reverse Cholesterol Transport. *Circ. Res.* **2019**, *124*, 1505–1518. [[CrossRef](#)]
45. Karathanasis, S.K.; Freeman, L.A.; Gordon, S.M.; Remaley, A.T. The Changing Face of HDL and the Best Way to Measure It. *Clin. Chem.* **2017**, *63*, 196–210. [[CrossRef](#)]
46. Buring, J.E.; O’Connor, G.T.; Goldhaber, S.Z.; Rosner, B.; Herbert, P.N.; Blum, C.B.; Breslow, J.L.; Hennekens, C.H. Decreased HDL2 and HDL3 cholesterol, Apo A-I and Apo A-II, and increased risk of myocardial infarction. *Circulation* **1992**, *85*, 22–29. [[CrossRef](#)]
47. Salonen, J.T.; Salonen, R.; Seppänen, K.; Rauramaa, R.; Tuomilehto, J. HDL, HDL2, and HDL3 subfractions, and the risk of acute myocardial infarction. A prospective population study in eastern Finnish men. *Circulation* **1991**, *84*, 129–139. [[CrossRef](#)]
48. Deng, X.; Ma, J.; Song, M.; Jin, Y.; Ji, C.; Ge, W.; Guo, C. Effects of products designed to modulate the gut microbiota on hyperlipidaemia. *Eur. J. Nutr.* **2019**, *58*, 2713–2729. [[CrossRef](#)]
49. Zheng, M.M.; Wang, R.F.; Li, C.X.; Xu, J.H. Two-step enzymatic synthesis of ursodeoxycholic acid with a new 7β-hydroxysteroid dehydrogenase from Ruminococcus torques. *Process Biochem.* **2015**, *50*, 598–604. [[CrossRef](#)]
50. IOC consensus statement on sports nutrition 2010. *J. Sports Sci.* **2011**, *29*, S3–S4. [[CrossRef](#)]
51. Joo, J.; Williamson, S.A.; Vazquez, A.I.; Fernandez, J.R.; Bray, M.S. The influence of 15-week exercise training on dietary patterns among young adults. *Int. J. Obes.* **2019**, *43*, 1681–1690. [[CrossRef](#)]

52. Kujala, U.M.; Palviainen, T.; Pesonen, P.; Waller, K.; Sillanpää, E.; Niemelä, M.; Kangas, M.; Vähä-Ypyä, H.; Sievänen, H.; Korpelainen, R.; et al. Polygenic Risk Scores and Physical Activity. *Med. Sci. Sport. Exerc.* **2020**, *52*, 1518–1524. [[CrossRef](#)]
53. Sillanpää, E.; Palviainen, T.; Ripatti, S.; Kujala, U.; Kaprio, J. Polygenic Score for Physical Activity Is Associated with Multiple Common Diseases. *Med. Sci. Sport. Exerc.* **2021**, *53*, 193–194. [[CrossRef](#)]
54. Grieneisen, L.; Dasari, M.; Gould, T.J.; Björk, J.R.; Grenier, J.-C.; Yotova, V.; Jansen, D.; Gottel, N.; Gordon, J.B.; Learn, N.H.; et al. Gut microbiome heritability is nearly universal but environmentally contingent. *Science* **2021**, *373*, 181–186. [[CrossRef](#)]
55. Neuberger-Castillo, L.; Hamot, G.; Marchese, M.; Sanchez, I.; Ammerlaan, W.; Betsou, F. Method Validation for Extraction of DNA from Human Stool Samples for Downstream Microbiome Analysis. *Biopreserv. Biobank.* **2020**, *18*, 102–116. [[CrossRef](#)]
56. Rintala, A.; Riikonen, I.; Toivonen, A.; Pietilä, S.; Munukka, E.; Pursiheimo, J.P.; Elo, L.L.; Arikoski, P.; Luopajarvi, K.; Schwab, U.; et al. Early fecal microbiota composition in children who later develop celiac disease and associated autoimmunity. *Scand. J. Gastroenterol.* **2018**, *53*, 403–409. [[CrossRef](#)]
57. Bokulich, N.A.; Subramanian, S.; Faith, J.J.; Gevers, D.; Gordon, J.I.; Knight, R.; Mills, D.A.; Caporaso, J.G. Quality-filtering vastly improves diversity estimates from Illumina amplicon sequencing. *Nat. Methods* **2013**, *10*, 57–59. [[CrossRef](#)]
58. Segata, N.; Izard, J.; Waldron, L.; Gevers, D.; Miropolsky, L.; Garrett, W.S.; Huttenhower, C. Metagenomic biomarker discovery and explanation. *Genome Biol.* **2011**, *12*, 1–18. [[CrossRef](#)]
59. Soininen, P.; Kangas, A.J.; Würtz, P.; Suna, T.; Ala-Korpela, M. Quantitative Serum Nuclear Magnetic Resonance Metabolomics in Cardiovascular Epidemiology and Genetics. *Circ. Cardiovasc. Genet.* **2015**, *8*, 192–206. [[CrossRef](#)]
60. Soininen, P.; Kangas, A.J.; Würtz, P.; Tukiainen, T.; Tynkkynen, T.; Laatikainen, R.; Järvelin, M.-R.; Kähönen, M.; Lehtimäki, T.; Viikari, J.; et al. High-throughput serum NMR metabolomics for cost-effective holistic studies on systemic metabolism. *Analyst* **2009**, *134*, 1781–1785. [[CrossRef](#)]