



**HAL**  
open science

# Understanding the Holobiont: Crosstalk Between Gut Microbiota and Mitochondria During Long Exercise in Horse

Núria Mach, Marco Moroldo, Andrea Rau, Jérôme Lecardonnel, Laurence Le Moyec, Céline Robert, Eric Barrey

► **To cite this version:**

Núria Mach, Marco Moroldo, Andrea Rau, Jérôme Lecardonnel, Laurence Le Moyec, et al.. Understanding the Holobiont: Crosstalk Between Gut Microbiota and Mitochondria During Long Exercise in Horse. *Frontiers in Molecular Biosciences*, 2021, 8, 10.3389/fmolb.2021.656204 . hal-03197210

**HAL Id: hal-03197210**

**<https://hal.inrae.fr/hal-03197210v1>**

Submitted on 13 Apr 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



# Understanding the Holobiont: Crosstalk Between Gut Microbiota and Mitochondria During Long Exercise in Horse

Núria Mach<sup>1\*</sup>, Marco Moroldo<sup>1</sup>, Andrea Rau<sup>1,2</sup>, Jérôme Lecardonnel<sup>1</sup>, Laurence Le Moyec<sup>3,4</sup>, Céline Robert<sup>1,5</sup> and Eric Barrey<sup>1</sup>

<sup>1</sup>Université Paris-Saclay, INRAE, AgroParisTech, GABI, 78350, Jouy-en-Josas, France, <sup>2</sup>BioEcoAgro Joint Research Unit, INRAE, Université de Liège, Université de Lille, Université de Picardie Jules Verne, Estrées-Mons, France, <sup>3</sup>Université d'Évry Val d'Essonne, Université Paris-Saclay, Évry, France ABI UMR 1313, INRAE, Université Paris-Saclay, AgroParisTech, Jouy-en-Josas, France, <sup>4</sup>MCAM UMR7245, CNRS, Muséum National d'Histoire Naturelle, Paris, France, <sup>5</sup>École Nationale Vétérinaire d'Alfort, Maisons-Alfort, France

## OPEN ACCESS

### Edited by:

Amelia Palermo,  
University of California, Los Angeles,  
United States

### Reviewed by:

Tao Huan,  
University of British Columbia, Canada  
Fumio Matsuda,  
Osaka University, Japan

### \*Correspondence:

Núria Mach  
nuria.mach@inrae.fr

### Specialty section:

This article was submitted to  
Metabolomics,  
a section of the journal  
Frontiers in Molecular Biosciences

Received: 20 January 2021

Accepted: 16 February 2021

Published: 08 April 2021

### Citation:

Mach N, Moroldo M, Rau A, Lecardonnel J, Le Moyec L, Robert C and Barrey E (2021) Understanding the Holobiont: Crosstalk Between Gut Microbiota and Mitochondria During Long Exercise in Horse. *Front. Mol. Biosci.* 8:656204. doi: 10.3389/fmolb.2021.656204

Endurance exercise has a dramatic impact on the functionality of mitochondria and on the composition of the intestinal microbiome, but the mechanisms regulating the crosstalk between these two components are still largely unknown. Here, we sampled 20 elite horses before and after an endurance race and used blood transcriptome, blood metabolome and fecal microbiome to describe the gut-mitochondria crosstalk. A subset of mitochondria-related differentially expressed genes involved in pathways such as energy metabolism, oxidative stress and inflammation was discovered and then shown to be associated with butyrate-producing bacteria of the *Lachnospiraceae* family, especially *Eubacterium*. The mechanisms involved were not fully understood, but through the action of their metabolites likely acted on *PPAR $\gamma$* , the *FRX-CREB* axis and their downstream targets to delay the onset of hypoglycemia, inflammation and extend running time. Our results also suggested that circulating free fatty acids may act not merely as fuel but drive mitochondrial inflammatory responses triggered by the translocation of gut bacterial polysaccharides following endurance. Targeting the gut-mitochondria axis therefore appears to be a potential strategy to enhance athletic performance.

**Keywords:** endurance, holobiont, horse, metabolome, microbiota, mitochondria, transcriptome

## INTRODUCTION

To keep up with energy demand and to maintain homeostasis, endurance exercise modifies multiple systems, ranging from the whole-body level to the molecular level (Clark and Mach, 2017; Mach and Fuster-Botella, 2017). In recent years, our understanding of the role played by mitochondria during this kind of challenge has expanded far beyond its bioenergetic capacity, which is represented by well characterized pathways such as oxidative phosphorylation (OXPHOS), fatty acid  $\beta$ -oxidation (FAO) and the tricarboxylic acid (TCA) cycle (Pfanner et al., 2019). Indeed, it is now widely accepted that mitochondria regulate cytosolic calcium homeostasis and cellular redox status, that they generate much of the cell reactive oxygen species (ROS), and that they are involved in steroid and heme biosynthesis, urea degradation, apoptosis and initiation of inflammation through inflammasomes (Chinnery and Hudson, 2013; Wong et al., 2016; Jackson and Theiss, 2020; Vezza et al., 2020).

It has also been established that the mitochondrial genome (mtDNA) and the nuclear genome are constantly communicating with each other to regulate the aforementioned pathways. For example, most of the proteins involved in OXPHOS and mitochondrial functions like mtDNA replication and expression, mtDNA repair, redox and energy regulation are encoded by the nuclear genome and require specific targeting signals to be directed from the cytosol to mitochondrial surface receptors and then to the proper mitochondrial sub-compartments (Pfanter et al., 2019). The transcriptional programs of mitochondria comprise over 1,600 nuclear-encoded mitochondrial proteins (Pagliarini et al., 2008; Richter-Dennerlein et al., 2016). Any alteration in the OXPHOS and FAO processes, in mitochondrial membrane potential ( $\Delta\Psi_m$ ), mitochondrial biogenesis, ROS production and inflammation can have a deep impact on the response to endurance exercise. For instance, increased mitochondrial biogenesis improves muscle endurance performance due to higher rates of OXPHOS and FAO (Hood et al., 2011). On the contrary, lower metabolic rates, increased ROS production and acidosis during prolonged exercise are associated with fatigue and inability to maintain speed (Rapoport, 2010). Endurance or long exercise and mitochondrial functions are strictly intertwined and influence each other.

The gut microbiota is considered a central organ because of its direct and indirect roles in host physiology, including improved metabolic health and athletic performance (Barton et al., 2017; Keohane et al., 2019; Scheiman et al., 2019). A healthy ecosystem in horse includes numerous, highly dominant taxa along with a multitude of minor players with lower representation, but important metabolic activity (Costa and Weese, 2018; Kauter et al., 2019; Mach et al., 2020).

The interdependence of gut microbiota and mitochondria is being increasingly recognized, with many diseases originating from mitochondrial dysfunctions linked to well-described changes in gut microbiota (Karlsson et al., 2013; Saint-Georges-Chaumet et al., 2015; Mottawea et al., 2016; Franco-Obregón and Gilbert, 2017; Bajpai et al., 2018; Cardoso and Empadinhas, 2018; Saint-Georges-Chaumet and Edeas, 2018; Yardeni et al., 2019). This complex interplay occurs principally through endocrine, immune and humoral signaling (Mottawea et al., 2016). Enteric short-chain fatty acids (SCFAs), the major class of metabolites produced from bacterial fermentation of non-digestible carbohydrates, are widely thought to mediate the relationship between the gut microbiota and the mitochondria in different tissues. Branched-chain amino acids (BCAAs), secondary bile acids, ROS, nitric oxide (NO), and hydrogen sulfide ( $H_2S$ ) are also thought to play at least a partial role in this molecular interchange (Clark and Mach, 2017).

While formal proof is missing in horse athletes, a prevailing hypothesis is that the gut microbiota and its metabolites regulate crucial transcription factors and coactivators involved in mitochondrial functions that underpin endurance performance (Hawley et al., 2018). In mice models, gut microbiota depletion via broad-spectrum antibiotics showed reduced production of SCFAs, lower bioavailability of serum glucose, decreased endurance capacity and impairment of the *ex vivo* skeletal muscle

contractile function (Nay et al., 2019). In close agreement, gut microbiota depletion also triggered a reduction of both faecal SCFA content and circulating concentration of SCFAs coupled to a drop in running capacity in mice (Okamoto et al., 2019). In contrast, mice with *Veillonella* in their intestinal ecosystem showed significantly increased submaximal treadmill run time to exhaustion (Scheiman et al., 2019), prompting the authors to speculate that the lactate generated during sustained bouts of exercise could be accessible to the microbiota and converted into SCFAs that ultimately enhanced energetic resilience and stamina. Alternatively, there may be other mechanisms through which gut microbiota and its metabolites relate to mitochondria, including but not limited to the regulation of mitochondrial oxidative stress (Jones and Neish, 2014; Franco-Obregón and Gilbert, 2017), as well as the activation of the inflammasome and the production of inflammatory cytokines, all of which are key players in the adaptation to endurance exercise (Clark and Mach, 2017; Mach and Fuster-Botella, 2017).

We recently presented a three-pronged association study, connecting horse gut microbiota with untargeted serum metabolome data and measures of host physiology and performance in the context of endurance (Plancade et al., 2019). We found no significant associations between the gut ecosystem and serum metabolites, especially those relying heavily on mitochondrial OXPHOS, FAO, TCA and gluconeogenesis. The number of annotated metabolites in the study was likely insufficient to reliably encompass the given mitochondrial functions. To further advance our knowledge of the molecular basis for the gut-mitochondria crosstalk that support the adaptation to long exercise, in this work we tethered whole blood transcriptome profiling to our previous metabolome and metagenome data. In doing so, we sought to identify the ways in which mitochondrial and nuclear transcriptomes coordinate with each other, and how gut microbiota and circulating metabolites can dynamically modulate this process. By jointly characterizing the whole blood transcriptome, metabolome, fecal microbiota and SCFAs of 20 elite horses competing in an endurance race, we aim to provide a functional readout of microbial activity and improve our understanding of the gut microbiota-mitochondria axis during long exercise.

## MATERIALS AND METHODS

### Ethics Approval

The study protocol was reviewed and approved by the local animal care and use committee (ComEth EnvA-Upec-ANSES, reference: 11-0041, dated July 12th, 2011) for horse study. All the protocols were conducted in accordance with EEC regulation (n° 2010/63/UE) governing the care and use of laboratory animals, which has been effective in France since the 1st of January 2013. In all cases, the owners and riders provided their informed consent prior to the start of sampling procedures with the animals.

### Animals

Twenty pure-breed or half-breed Arabian horses (7 females, 3 male, and 10 geldings; mean  $\pm$  SD age:  $10 \pm 1.69$ ) were selected

from the cohort used by Placade et al. (2019) (**Supplementary Table S1**). The 20 horses were selected following these criteria: 1) enrollment in the 160 km or 120 km category; 2) blood sample collection before and after the race; 3) feces collection before the race; 4) absence of gastrointestinal disorders during the four months prior to enrollment; 5) absence of antibiotic treatment during the four months prior to enrollment and absence of anthelmintic medication within 60 days before the race; 6) a complete questionnaire about diet composition and intake.

Among the 20 horses selected for this study, 16 horses were enrolled for the 160 km category and four for the 120 km category. In the 160 km category, two animals were eliminated due to tiredness after 94 km and 117 km, respectively, and five horses failed a veterinary gate check due to lameness after 94 km ( $n = 1$ ), and after 117 km ( $n = 4$ ). In the 120 km category, one horse was eliminated due to metabolic troubles after 90 km (**Supplementary Table S1**).

The weather conditions, terrain difficulty and altitude were the same for all the participants enrolled in the study as all races (120 and 160 km) took place during October 2015 in Fontainebleau (France). The average air temperature was 15°C, with a maximum of 20°C and a minimum of 11°C, the average air humidity was 88%, and no rain was recorded.

As detailed by Placade et al. (2019), to ensure sample homogeneity, the participating horses were subject to the same management practices throughout the endurance ride and passed the International Equestrian Federation (FEI) compulsory examination before the start. Animals were fed 2–3 h before the start of the endurance competition with *ad libitum* hay and 1 kg of concentrate pellets. During the endurance competition, all the animals underwent veterinary checks every 20–40 km, followed by recovery periods of 40–50 min (in accordance with the FEI endurance rules). After each veterinary gate check, the animals were provided with *ad libitum* water and hay and a small amount of concentrate pellets.

## Transcriptomic Microarray Data Production, Pre-processing and Analysis

Blood samples for RNA extraction were collected from each animal at T0 and T1 using Tempus Blood RNA tubes (Thermo Fisher). Because blood interacts with every organ and tissue in the body and has crucial roles in immune response, inflammation and physiological homeostasis (Mohr and Liew, 2007), blood-based transcriptome was carried-out as a means for exploring the response to endurance.

Total RNAs were then isolated using the Preserved Blood RNA Purification Kit I (Norgen Biotek Corp., Ontario, Canada), according to the manufacturer's instructions. RNA purity and concentration were determined using a NanoDrop ND-1000 spectrophotometer (Thermo Fisher) and RNA integrity was assessed using a Bioanalyzer 2100 (Agilent Technologies, Santa Clara, CA, United States). All the 40 RNA samples were processed.

Transcriptome profiling was performed using an Agilent 4 × 44K horse custom microarray (Agilent Technologies,

AMADID 044466). All of the steps were performed by the @BRIDGe facility (INRAE Jouy-en-Josas, France, <http://abridge.inra.fr/>), as described previously (Mach et al., 2016; 2017b).

The horse array was annotated as described by Mach et al. (2016), Mach et al. (2017b). In a limited number of cases, a manual annotation step was also included. Probe intensities were background-corrected using the “normexp” function, log<sub>2</sub> scaled and quantile normalized using the limma package (version 3.1.42.2) (Smyth, 2004) in the R environment (version 3.6.1). Quantile normalization seeks to reduce technical variability and impose identical empirical distributions between samples to facilitate comparisons (Smyth, 2004). Only the probes which presented, on at least two arrays, intensity values at least 10% higher than the 95% percentile of all the negative control probes were kept. Subsequently, controls were discarded and the probes corresponding to genes were summarized. The obtained expression matrix “E1” was processed with the arrayQualityMetrics R package (version 3.42.0) (Kauffmann et al., 2009) for quality assessment. No outliers were detected. The resulting pre-processed normalized expression values were thus approximately normally distributed and suited for differential analysis via conventional linear models.

The differential analysis was performed using the limma R package. A linear model was fitted for each gene, setting the time, the sex, the distance and whether the animal was eliminated from the race as fixed effects, and comparing T1 to T0. The individual was included as a random effect using the “duplicateCorrelation” function (Smyth et al., 2005). The *p*-values were Bonferroni corrected setting a threshold of 0.05. The expression matrix “E1” was then used to perform a scaled principal component analysis (PCA) with FactoMineR R package (version 2.4) (Lé et al., 2008).

To confirm the results of the DE analysis, the weighted gene co-expression network analysis (WGCNA) method was also run on the “E1” matrix using the WGCNA R package (version 1.69) (Langfelder and Horvath, 2008). The parameters for the analysis were set as follows: “corFnc” = bicor, “type” = signed hybrid, “beta” = 10, “deepSplit” = 4, “minClusterSize” = 30, and “cutHeight” = 0.1. The eigengenes corresponding to each identified module were correlated individually to all the <sup>1</sup>H NMR and biochemical assay metabolites, i.e., a set of 56 different molecules (see next paragraphs). A module was then considered positively or negatively associated to this set of molecules if the Pearson *r* correlation values were  $\geq |0.65|$  for at least 5 molecules and if all the corresponding *p*-values were  $\leq 1e-05$ . The positively and the negatively correlated modules defined in this way were merged to obtain a single gene list, which was subsequently compared to the differentially expressed genes (DEGs) list using a Venn diagram.

Afterward, a literature-based meta-analytical enrichment test was carried out to assess whether the DEG list and the gene list obtained using WGCNA were enriched in genes related to mitochondria. To this aim, first a consensus list of genes related to these organelles was created (**Supplementary Information and Supplementary Table S2**). This gene list was then annotated in a descriptive way. First, gene symbols were converted into the corresponding KEGG Orthology (KO)

codes using the “db2db” tool (<https://biodbnet-abcc.ncifcrf.gov/db/db2db.php>) from the bioDBnet suite and using the then up-to-date underlying databases. Then, the retrieved KO codes were processed with the “Reconstruct Pathway” tool ([https://www.genome.jp/kegg/tool/map\\_pathway.html](https://www.genome.jp/kegg/tool/map_pathway.html)) from the KEGG Mapper suite. Eventually, the obtained KEGG pathways underwent some manual editing step to make them easier to interpret, namely 1) only first- and third-level hierarchies including at least 15 genes were kept for visualization; 2) the “Human diseases” first-level hierarchy, with all its child taxonomic terms, was removed; 3) the “Metabolic pathways” third-level hierarchy was discarded as it was redundant with respect to the other third-level hierarchies of the “Metabolism” term.

Then, the same gene list was intersected with the genes found expressed on the microarray (i.e., the “E1” matrix). The subset thus obtained was separately intersected with the DEG list and the WGCNA gene modules. A Fisher’s exact test and hypergeometric test were then used to evaluate the overrepresentation of genes related to mitochondrial functions in each list.

The intersection between the genes related to mitochondria found on the microarray and the DEGs was referred to as “mt-related genes”. It was functionally annotated using ClueGO (version 2.5.7) (Bindea et al., 2009) by carrying out a right-sided test. Significance was set at a Benjamini–Hochberg adjusted *p*-value of 0.05 and the *k*-score was fixed at 0.4. Only the “KEGG January 30, 2019” ontology was selected.

## Proton Magnetic Resonance (<sup>1</sup>H NMR) Metabolite Analysis in Plasma

As described by Plancade et al. (2019) and Le Moyec et al. (2019) to characterize the metabolic phenotype of endurance horses in detail, we measured <sup>1</sup>H NMR spectra at 600 MHz for plasma samples. Blood was collected from each horse the day before the event (T0) and within 30 min from the end of the endurance race (T1) using sodium fluoride and oxalate tubes in order to inhibit further glycolysis that may increase lactate levels after sampling. All the samples were immediately refrigerated at 4°C to minimize the metabolic activity of cells and enzymes and to keep metabolite composition as stable as possible, and clotting time was strictly controlled to avoid cell lysis. After clotting, plasma was separated from blood cells and subsequently transported to the laboratory at 4°C and then frozen at –80°C (no more than 5 h later in all cases). Plasma samples were subsequently thawed at room temperature. Using 5 mm NMR tubes, 600 μL of plasma were added to 100 μL deuterium oxide for field locking. The <sup>1</sup>H NMR spectra were acquired at 500 MHz with an AVANCE III (Bruker, Billerica, MA, United States) equipped with a 5 mm reversed QXI Z-gradient high-resolution probe. Water signal was suppressed with a pre-saturation pulse ( $3.42 \times 10^{-5}$  W) during a 3s-relaxation delay at the water resonance frequency. The spectrum was divided into 0.001 ppm regions (bins) over which the signals were integrated to obtain intensities. The high- and low-field ends were removed, leaving only the data between 9.5 and 0.0 ppm. The region between 4.5 and 5.0 ppm, which corresponded to the signal of residual water, was also removed. Chemical signals were

first normalized according to a relevant internal standard, the so-called C1- $\alpha$ -glucose doublet, with signal at 5.23 ppm. Next, the bins (which exhibit skewed distributions across samples) were normalized according to the spectra using the probabilistic quotient method to correct for unwanted technical biases, and were subsequently centered and scaled to unit variance (Dieterle et al., 2006). Further details on sample preparation, data acquisition, data quality control, spectroscopic data pre-processing, and data pre-processing including bin alignment, normalization, scaling and centering are broadly discussed elsewhere (Le Moyec et al., 2014).

As specified in Plancade et al. (2019), metabolite identification was then performed by using information acquired from other available biochemical databases, namely HMD (<http://www.hmdb.ca/>), KEGG (<https://www.genome.jp/kegg/>), METLIN (<http://metlin.scripps.edu/>), ChEBI (<http://www.ebi.ac.uk/Databases/>), and LIPID MAPS (<http://www.lipidmaps.org/>) and the literature (Le Moyec et al., 2014; Le Moyec et al., 2019; Mach et al., 2017b; Jang et al., 2017). Each peak was assigned to a metabolite when chemical shifts of peaks in the samples were the same as in the publicly available reference databases or literature (with a shift tolerance level of  $\pm 0.005$  ppm), in order to counteract the effects of measurements and pre-processing variability introduced by factors such as pH and solvents. A manual curation for identified compounds was carried out by an expert in horse metabolomics (Le Moyec et al., 2014). Eventually, the relative abundance of each metabolite was calculated as the area under the peak (Zheng et al., 2011). A total of 50 metabolites was identified, which belonged to the following broad categories: amino acids, including aromatic and branched-chain amino acids, energy metabolism-related metabolites, saccharides, and organic osmolytes (**Supplementary Table S3**). We refer to our previous work (Plancade et al., 2019) for more detailed descriptions of the pre-processing and main results of the plasma metabolome data that were used to generate the input files provided with this study.

## Biochemical Assay Data Production

Blood samples for biochemical assays were collected at T0 and T1 using 10 ml BD Vacutainer EDTA tubes (Becton Dickinson, Franklin Lakes, NJ, United States). As detailed in Plancade et al. (2019), after clotting the tubes were centrifuged and the harvested serum was stored at 4°C until analysis (no more than 48 h later, in all cases). Sera were assayed for total bilirubin, conjugated bilirubin, total protein, creatinine, creatine kinase,  $\beta$ -hydroxybutyrate, aspartate transaminase (ASAT),  $\gamma$ -glutamyltransferase and serum amyloid A levels on a RX Imola analyzer (Randox, Crumlin, United Kingdom). The biochemical values obtained are reported in the **Supplementary Table S4**.

## Fecal Measurements, 16S Data Production and Analysis

Fresh fecal samples were obtained while monitoring the horses before the race (no more than 24 h before starting the race, in all cases). One fecal sample from each animal was collected off the

ground immediately after defecation as described by Mach et al. (2017a) and Plancade et al. (2019), and three aliquots (200 mg) were prepared. Since most of the horses experienced dehydration after the race, the gastrointestinal emptying was significantly delayed and therefore it was not possible to recover the feces after the race.

Aliquots for pH determination were kept at room temperature, while aliquots for SCFA analysis and DNA extraction were snap-frozen. Since most of the horses experienced dehydration after the race, the gastrointestinal emptying was significantly delayed and consequently we were not able to recover the feces after the race.

Fecal pH was immediately determined after 10% fecal suspension (wt/vol) in saline solution (0.15 M NaCl solution). SCFAs concentrations were measured as previously described in Mach et al. (2017a). The values obtained are described in the **Supplementary Table S5**.

Total DNA was extracted using the EZNA Stool DNA Kit (Omega Bio-Tek, Norcross, Georgia, United States), and following the manufacturer's instructions. DNA was then quantified using a Qubit and a dsDNA HS assay kit (Thermo Fisher).

The V3-V4 hyper-variable region of the 16S rRNA gene was amplified as previously reported by our team (Mach et al., 2017a, Mach et al., 2020; Clark et al., 2018; Massacci et al., 2019; Plancade et al., 2019). The concentration of the purified amplicons was measured using a Nanodrop 8000 spectrophotometer (Thermo Fisher) and their quality was checked using DNA 7500 chips onto a Bioanalyzer 2100 (Agilent Technologies). All libraries were pooled at equimolar concentration, and the final pool had a diluted concentration of 5 nM and was used for sequencing. The pooled libraries were mixed with 15% PhiX control according to the protocol provided by Illumina (Illumina, San Diego, CA, United States) and sequenced on a single MiSeq (Illumina, United States) run using a MiSeq Reagent Kit v2 (500 cycles).

The Divisive Amplicon Denoising Algorithm (DADA) was implemented using the DADA2 plug-in for QIIME 2 (version 2019.10) to perform quality filtering and chimera removal and to construct a feature table consisting of read abundance per amplicon sequence variant (ASV) by sample (Callahan et al., 2016). DADA2 models the amplicon sequencing error in order to identify unique ASV and infers sample composition more accurately than traditional Operational Taxonomic Unit (OTU) picking methods that identify representative sequences from clusters of sequences based on a % similarity cut-off (Callahan et al., 2016). The output of DADA2 was an abundance table, in which each unique sequence was characterized by its abundance in each sample. Taxonomic assignments were given to ASVs by importing SILVA 16S representative sequences and consensus taxonomy (release 132, 99% of identity) to QIIME 2 and classifying representative ASVs using the naive Bayes classifier plug-in (Bokulich et al., 2018). The feature table, taxonomy, and phylogenetic tree were then exported from QIIME 2 to the R statistical environment and combined into a phyloseq object (McMurdie and Holmes, 2013). Prevalence filtering was

applied to remove ASVs with less than 1% prevalence and in fewer than three individuals, decreasing the possibility of data artifacts affecting the analysis (Callahan et al., 2016). To reduce the effects of uncertainty in ASV taxonomic classification, we conducted most of our analysis at the microbial genus level.

The phyloseq (version 1.32.0) (McMurdie and Holmes, 2012), vegan (version 2.5.6) (Dixon, 2003) and microbiome packages (version 1.10.0) were used in R (version 4.0.2) for the downstream steps of analysis. The minimum sampling depth in our data set was 10,423 reads per sample. Reads were clustered into 3,385 chimera- and singleton-filtered Amplicon Sequence variants (ASVs) at 99% sequence similarity (**Supplementary Data Sheet 1**). ASV counts per sample and ASV taxonomical assignments are available in **Supplementary Table S6**. Abundance data were aggregated at genus, family, order, class and phyla levels throughout the taxonomic-agglomeration method in the phyloseq R package, which merges taxa of the same taxonomic category for a user-specific taxonomic level. The genera abundance data were scaled to proportions to correct the differences in read depth. This processing step was performed by scaling the reads for each taxon in a given sample by the total number of reads in that sample. The resulting table of proportions included 100 genera (**Supplementary Table S8**).

## qPCR Quantification of Bacterial, Fungal and Protozoan Concentration

As detailed by Plancade et al. (2019), concentrations of bacteria, anaerobic fungi and protozoa in fecal samples were quantified by qPCR using a QuantStudio 12K Flex platform (Thermo Fisher Scientific, Waltham, United States). Primers for real-time amplification of bacteria (FOR: 5'-CAGCMGCCGCGGTAANWC-3'; REV: 5'-CCGTCAATTCMTTTRAGTTT-3'), anaerobic fungi (FOR: 5'-TCCTACCCTTTGTGAATTTG-3'; REV: 5'-CTGCGT TCTTCATCGTTGCG-3') and protozoa (FOR: 5'-GCTTTCGWGTGGTAGTGATT-3'; REV: 5'-CTTGCCCTCYAATCGTWCT-3'), are described in Mach et al. (2015) and Clark et al. (2018) and were purchased from Eurofins Genomics (Ebersberg, Germany).

Amplified fragments of the target amplicons were used to create a seven-point 10-fold standard dilution series. The dilution points ranged from  $2.25 \times 10^7$  to  $2.25 \times 10^{13}$  copies per  $\mu\text{g}$  of DNA for bacteria and protozoa and from  $3.70 \times 10^6$  to  $3.70 \times 10^{12}$  copies per  $\mu\text{g}$  of DNA for anaerobic fungi. qPCR reactions were performed in a final volume of 20  $\mu\text{L}$ , containing 10  $\mu\text{L}$  of Power SYBR Green PCR Master Mix (Thermo Fisher), 2  $\mu\text{L}$  of standard or DNA template at 0.5 ng/ $\mu\text{L}$  and 0.6  $\mu\text{M}$  of each primer to a final concentration of 200 mM for bacteria and anaerobic fungi and 150 mM for protozoa.

In all the cases, the thermal cycling conditions were as follows: initial denaturation at 95 °C for 10 min; 40 cycles of denaturation at 95°C for 15 s, annealing and extension at 60°C for 60 s. To check for the absence of nonspecific signals, a dissociation step was added after each amplification. It was carried out by ramping the temperature from 60 to 95°C. All qPCR runs were performed in triplicate, and the standard curve obtained using the target

amplicons was used to calculate the number of copies of microorganisms in feces.

Taking into account the molecular mass of nucleotides and the amplicon length, the number of copies was obtained using the following equation: copies per nanogram =  $(NL \times A \times 10^{-9}) / (n \times mw)$ , where “NL” is the Avogadro constant ( $6.02 \times 10^{23}$  molecules per mole), “A” is the molecular weight of DNA molecules (ng), “n” is the length of the amplicon in base pairs, and “mw” is the molecular weight per base pair.

The number of copies of bacteria, anaerobic fungi and protozoa were log transformed to get the sampled data in line with the assumptions of parametric statistics, i.e., that the residuals from a model fit are normally distributed with a homogeneous variance. Accordingly, the results are expressed as  $\log_{10}$  of gene copies (per gram of wet weight; **Supplementary Table S8**).

## Integrative Statistical Analysis

Data integration was carried out using several approaches and different combinations of data sets.

Prior to the integration, we applied an additional pre-processing step for the biochemical assay data, metabolome data and gene expression data. In particular, to eliminate intra-individual variability and focus on the respective differential signals between T1 and T0, we considered  $\Delta$  values (T1–T0) for each of these data sets, as previously described by Plancade et al. (2019). For the transcriptome, we constructed a matrix (“E2”) of log-transformed expression values between T1 and T0 (i.e., the difference in  $\log_2$ -normalized expression between T1 and T0, equivalent to the  $\log_2$  value of the T1/T0 ratio) for the differentially expressed mt-related genes.

In the case of fecal microbiota, since proportions data have been shown to perform poorly in differential abundance testing and data integration (Rohart et al., 2017), we have instead decided to apply a centered log-ratio (CLR) transformation to the genera count matrix using the mixMC framework of the mixOmics R package (version 6.10.9) (Rohart et al., 2017). The CLR transformation, which is well suited to highly sparse compositional data, consists in dividing each sample by the geometric mean of its values and calculating its logarithm. Raw data were first pre-filtered by removing the genera for which the percentage of the sum of counts was lower than 1% compared to the total sum of all counts. Then, the pre-filtered data were transformed using CLR and applying an offset of 1. The filtered genera matrix “G2” obtained in this way included 85 genera (**Supplementary Table S9**).

The integration of data was then performed using three different methods and working with all six data sets available, namely: 1)  $\Delta$  values of mt-related genes; 2)  $\Delta$  values of  $^1\text{H}$  NMR metabolites; 3)  $\Delta$  values of the biochemical assay metabolites; 4) the fecal SCFAs at T0; 5) the fecal 16S rRNA gene sequencing data at T0 (i.e., the “G2” matrix); and 6) the concentration of fecal microorganisms at T0.

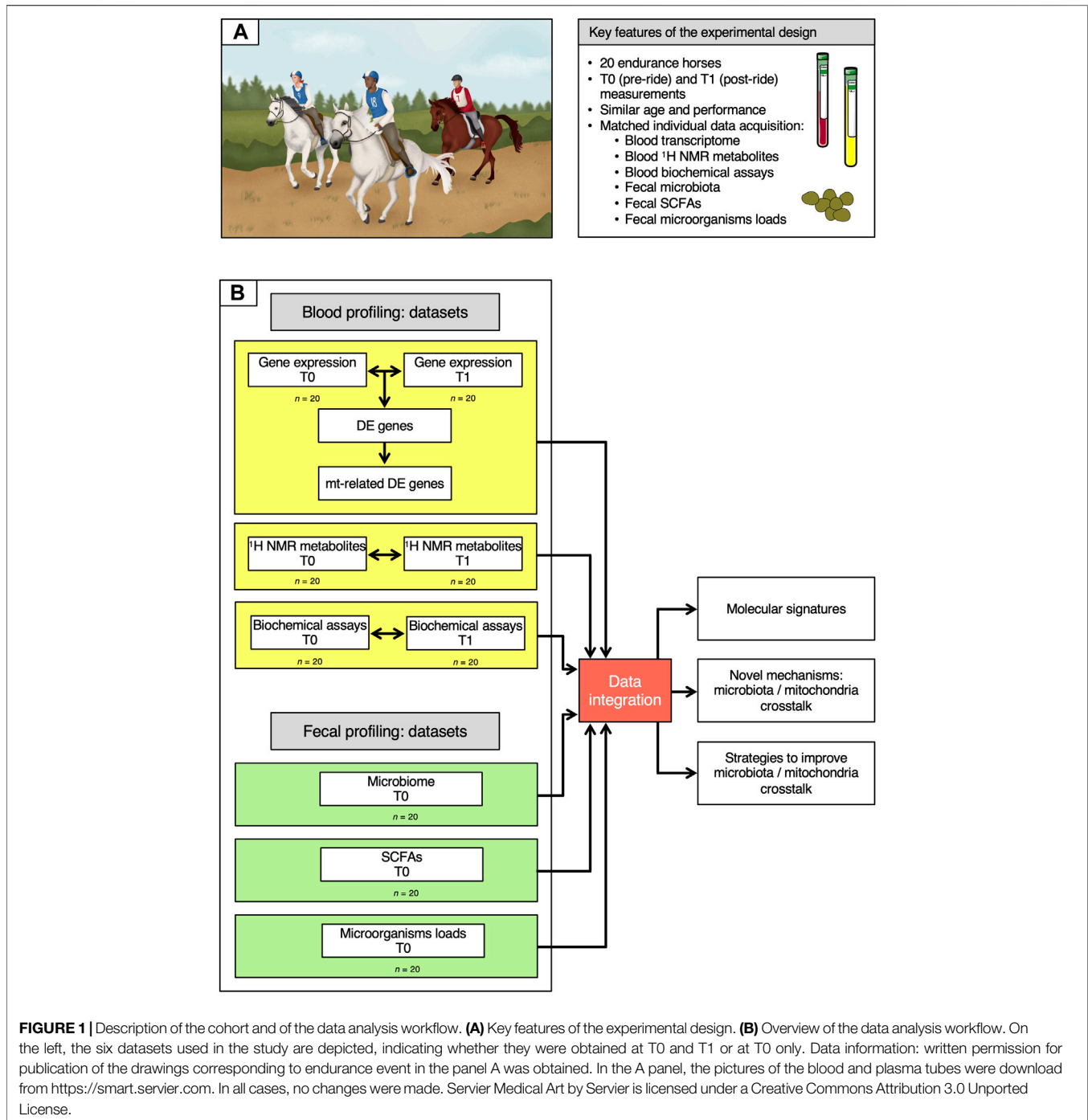
As a first integration approach, a global non-metric multidimensional scaling (NMDS) ordination was used to extract and summarize the variation in mt-related genes (the “response variable”) using the “metaMDS” function in vegan R

package. To determine the number of dimensions for each NMDS, stress values were calculated. Stress values are a measure of how much the distances in the reduced ordination space depart from the distances in the original p-dimensional space. High stress values indicate a greater possibility that the structuring of observations in the ordination space is entirely unrelated to that of the original full-dimensional space.

The other five data sets (the “explanatory variables”) were then fitted to the ordination plots using the “envfit” function in the vegan R package (Clarke and Ainsworth, 1993) with 10,000 permutations. The “envfit” function performs a multivariate analysis of variance (MANOVA) and linear correlations for categorical and continuous variables. The effect size and significance of each covariate were determined comparing the difference in the centroids of each group relative to the total variation, and all of the *p*-values derived from the “envfit” function were Benjamini–Hochberg adjusted. The obtained  $r^2$  gives the proportion of variability (that is, the main dimensions of the ordination) that can be attributed to the explanatory variables.

As a second integration approach, a forward-selection model-building method for redundancy analysis (RDA) (Blanchet et al., 2008) was used to extract and summarize the variation in mt-related genes (the “response variable”) that could be explained by the other five data sets (the “explanatory variables”). To determine which set of covariates provided the most parsimonious model, automatic stepwise model selection for constrained ordination methods was used as implemented by the “ordistep” function of the vegan R package. To test for robustness, a forward automatic model selection on a distance based RDA (db-RDA) was then performed using the “ordiR2step” function of the vegan package in R (Oksanen et al., 2013). This provided an estimation of the linear cumulative effect size of all the identified non-redundant covariates and of their independent fraction in the best model. In the case of this latter function, the mt-related genes matrix was modified using the Hellinger transformation prior to the analysis. These two RDA functions use different criteria for variable selection. The “ordistep” function uses the Akaike’s information criterion (AIC) and *p*-value < 0.05 obtained from Monte Carlo permutation tests, while “ordiR2step” uses the adjusted coefficient of determination ( $r_{adj}^2$ ). In both cases, the procedure begins by comparing a null model containing no variables and a test model containing one variable, where every possible covariate is considered.

As a third integrative approach, the N-integration algorithm DIABLO of the mixOmics R package was used. It is to be noted that, in the case of the N-integration algorithm DIABLO, the variables of all the data sets were also centered and scaled to unit variance prior to integration. In this case, the relationships existing among all six data sets were studied by adding a further categorical variable, i.e., the performance of horses. Horses that had a poor performance or that had been eliminated ( $n = 8$ ) were compared to horses that had completed the race ( $n = 12$ , **Supplementary Table S1**). DIABLO seeks to estimate latent components by modeling and maximizing the correlation between pairs of pre-specified datasets to unravel similar functional relationships between them

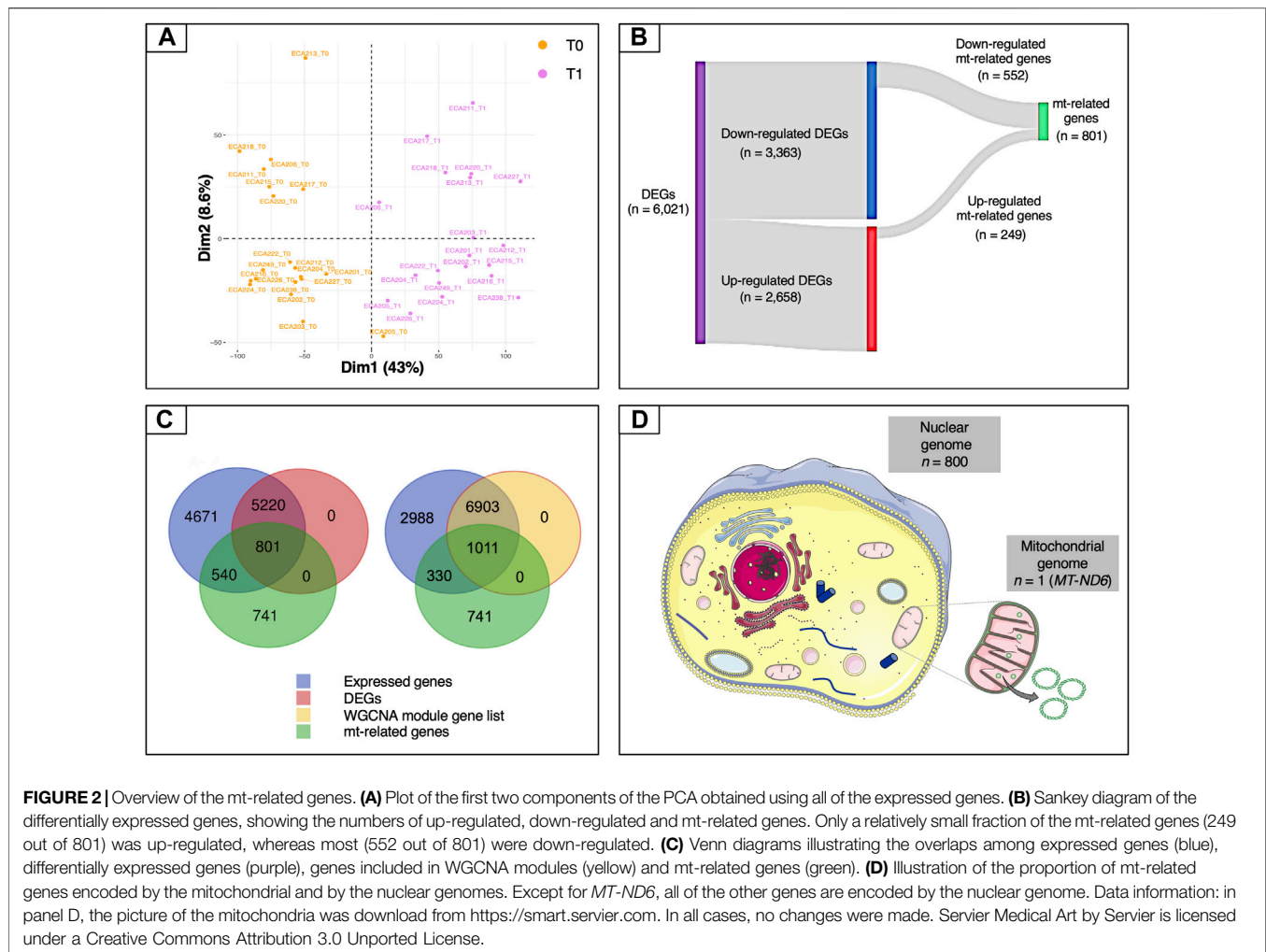


(Singh et al., 2019). A full weighted design was considered and, to predict the number of latent components and the number of discriminants, the “block.splsda” function was used. In both cases, the model was first fine-tuned using the leave-one-out cross-validation by splitting the data into training and testing. Then, classification error rates were calculated using balanced error rates (BERs) between the predicted latent variables with the centroid of the class labels (i.e., eliminated vs non eliminated horses) using the “max.dist” function. BERs account for

differences in the number of samples between different categories. Only interactions with  $r \geq |0.70|$  were visualized using CIRCOS. To visualize the high-confidence molecule co-associations determined by CIRCOS, only those with  $r \geq |0.70|$  and more than 15 connections were automatically visualized using the organic layout algorithm in Cytoscape (version 3.8.1) (<http://cytoscape.org>).

Finally, we performed a pairwise integration, focusing only on mt-related genes and microbiota data, and using the regularized





canonical correlation analysis (rCCA) method as implemented by the mixOmics R package (<https://cran.r-project.org/web/packages/mixOmics/>). The penalization parameters were estimated using the “shrinkage” method and setting the “ncomp” parameter to two. The correlation matrix thus obtained was filtered by retaining only the genes and the genera for which at least one association value data point presented  $r \geq |0.55|$ .

## RESULTS

### Cohort Characteristics

To elucidate how mitochondria and gut microbiota are linked during endurance exercise, we studied 20 healthy endurance horses selected from the cohort already described by PlanCADE et al. (2019) and Le Moyec et al. (2019) (Figure 1). All of the animals were of similar age and performance level (Supplementary Table S1). They performed a long exercise during about 8 h at an average speed of  $17.1 \pm 1.67$  km/h with some rest periods every 30–40 km.

Whole transcriptome profiles,  $^1\text{H}$  NMR metabolome profiles and biochemical assay data were obtained from blood samples

collected at both T0 (pre-ride) and T1 (post-ride), while SCFAs measurements, 16S rRNA data and the concentration of bacteria, anaerobic fungi and ciliate protozoa were generated from fecal samples at T0 alone.

### Blood Transcriptome Profiles and Mitochondrial Related Genes

To gather information about the global structure of the blood transcriptome, a scaled PCA was carried out on the expressed genes ( $n = 11,232$ ). The first component accounted for 43% of the total variability and revealed a marked separation of the two time points (Figure 2A). We then carried out a standard differential analysis between the two time points. After Bonferroni correction of the raw  $p$ -values, a total of 6,021 differentially expressed genes was obtained at an adjusted  $p < 0.05$ , of which 2,658 were upregulated and 3,363 downregulated at T1 respect to T0 (Supplementary Table S10; Figure 2B).

These results were then complemented using a WGCNA (Langfelder and Horvath, 2008) on the expressed genes. WGCNA identified three gene modules, corresponding to 7,914 genes, that were correlated to the  $^1\text{H}$  NMR and

biochemical assay metabolites (**Supplementary Table S11**). These genes strongly overlapped with the set of DEGs, 91.1% of which (i.e., 5,486 out of 6,021 genes) were included among them. The metabolites which showed the highest levels of correlation with the gene modules were bilirubin, non-esterified fatty acids (NEFAs), tyrosine, lactate and, to a lesser extent,  $\beta$ -hydroxybutyrate (BHB) (**Supplementary Figure S1**).

Because we were especially interested in understanding the role played by mitochondria in our biological system, we then decided to study in more detail the features related to these organelles. To this end, we used a literature-based meta-analytic approach to build a non-redundant consensus list of 2,082 genes related to mitochondria, based on the information available in the Integrated Mitochondrial Protein Index (IMPI) (Smith and Robinson, 2019), the Mitocarta Inventory (Calvo et al., 2016) and the literature (**Supplementary Information and Supplementary Table S2**). This consensus list was also descriptively annotated to gain global insight into the main biological functions represented within it. A total of 80 third-level KEGG hierarchies were identified, with a strong representation of pathways related to carbohydrate and lipid metabolism (8 and 9 ontology terms, respectively). Eight pathways were associated with amino acid metabolism, while five were linked to the apoptosis process (**Supplementary Figure S2**). This subset was crossed with each of the DEG and WGCNA module gene lists. In the case of the DEG list, the intersection with this subset yielded a total of 801 genes (**Supplementary Table S12 and Figure 2C**). Both the Fisher's exact test and the hypergeometric test showed strong levels of over-representation, with  $p$ -values of  $1.0 \times 10^{-5}$  and  $9.02 \times 10^{-7}$ , respectively. In the case of the WGCNA gene modules, the intersection included 1,011 genes (**Figure 2C**). Again, both statistical tests indicated strong enrichment, with  $p$ -values of  $1.0 \times 10^{-5}$  and  $1.07 \times 10^{-5}$ , respectively.

The set of 801 genes in the intersection of the mitochondrial consensus and DEG list, which will be referred to hereafter simply as "mt-related genes", was selected for the downstream steps of analysis. All of the mt-related genes were encoded by the nuclear genome except for *MT-ND6* (MT-NADH dehydrogenase, subunit 6), which was encoded by the antisense strand of the mt-DNA (**Figure 2D**). These mt-related genes were further characterized to gather information about their molecular function. The functional analysis showed that roughly 75% of these genes were directly involved in energy metabolism (i.e., pathways such as OXPHOS and FAO) and metabolite synthesis and degradation (**Supplementary Figure S3**). For instance, we observed an enrichment of genes related to nutrient transport across the mitochondrial inner membrane (*TOM/TIM* units, *VDAC*, *MPC1*, *ACAA1*, members of the mitochondrial carrier family *SLC25* and of the pyruvate dehydrogenase kinase isozyme), fatty acid metabolism (*ELOVL7*, *SIRT5*, and *ACAD* members), lipogenesis (*FASN* and *PPAR $\gamma$* ), and fatty acid channelling into oxidation (*CPT1B*, *ACADVL*, *ACOT9*, *ACOX1*, *ACSL1*, *ACSL4*, *ACSS3*). Additionally, key genes involved in the mitochondrial biogenesis (*POLG* and *POLG2*), mitochondrial fission (*PINK1*), mitochondrial fusion (*OPA1*), mitophagy (*BNIP3* and *PINK1*), oxidative stress (*SOD1*, *SOD2*,

glutathione S-transferases and glutathione peroxidases families), and the resolution of lipopolysaccharide induced pro-inflammatory pathway (*CIQBP*) were also found in this list. *CREB*, the most potent activator of *PGC-1 $\alpha$*  (Wu et al., 2006) was also differentially expressed upon endurance exercise.

Our data further indicated that among the mt-related genes, at least 21 genes encoding rps and rpl proteins of the small and large subunits of ribosomes (*rpl3*, *rpl4*, *rpl5*, *rpl6*, *rpl18*, *rpl23*, *rpl27*, *rpl36*, *rps3*, *rps8*, *rps9*, *rps10*, *rps11*, *rps12*, *rps13*, *rps14*, *rps15*, *rps16*, *rps17*, *rps18*, *rps19*) (Esser et al., 2004; Maier et al., 2013; Janouškovec et al., 2017) were common with the  $\alpha$ -Proteobacteria. This was also the case for the methionine sulfoxide reductase A (*MSRA*) and NAD(P)H dehydrogenase quinone 1 (*NQO1*) (Crisp et al., 2015), consistent with a vertical origin in the mitochondrial endosymbiont.

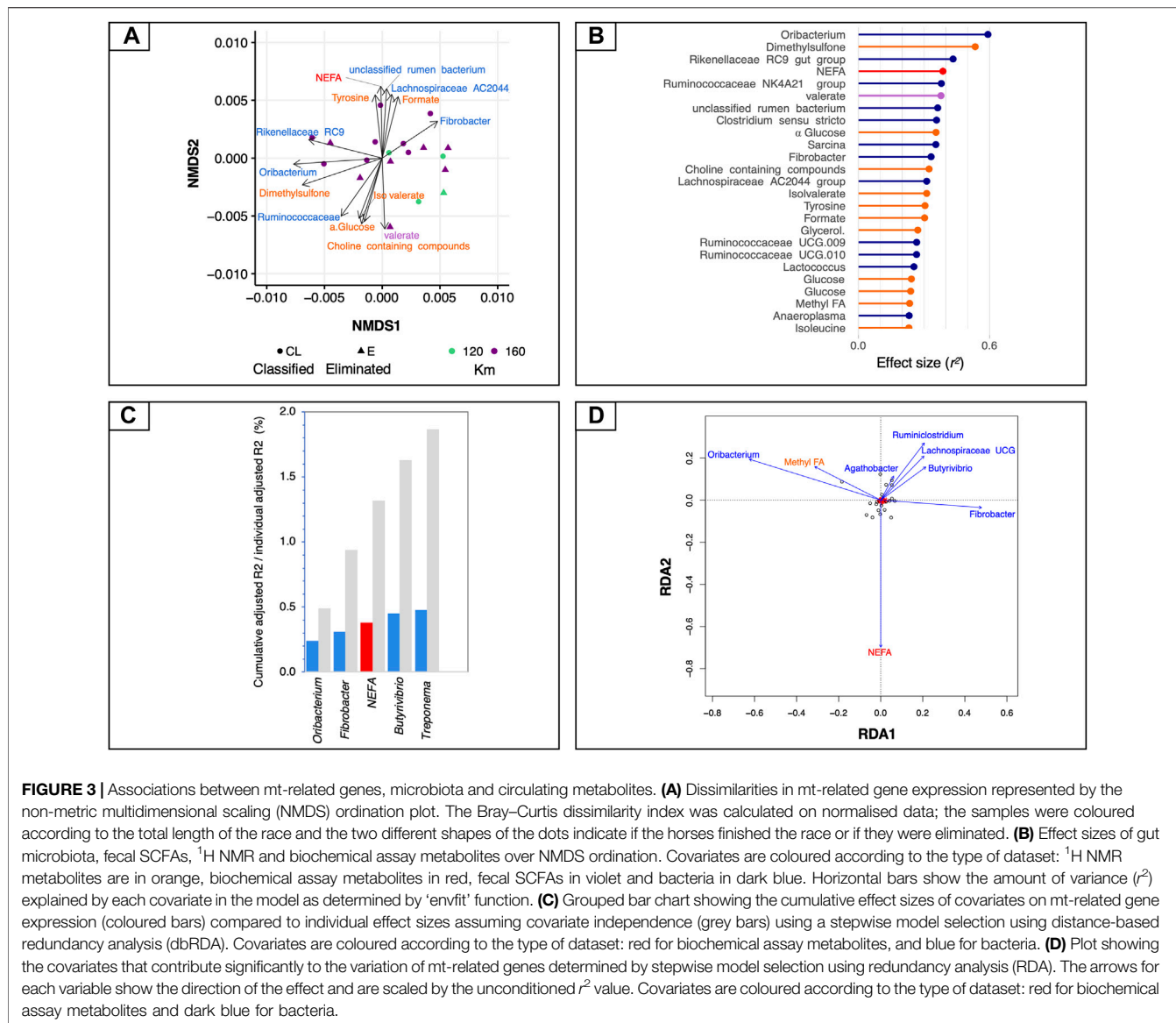
## <sup>1</sup>H NMR Metabolome, Biochemical Assay and Acetylcarnitine Profiles

The <sup>1</sup>H NMR metabolome and biochemical assay profiles used in this paper were gleaned from our previous works (Le Moyec et al., 2019; Plancade et al., 2019). Briefly, a total of 50 <sup>1</sup>H NMR known metabolites was detected in the plasma, including several amino acids, energy metabolism-related metabolites and organic osmolytes (**Supplementary Table S3**). Three well-known microbial derived metabolites were ascertained, including formate, dimethyl sulfone and trimethyl amine oxide (TMAO).

The relative abundance values of these circulating metabolites fell within the normal reference range for healthy horses. However, the concentration of lactate (a proxy for glycolytic stress and disturbances in cellular homeostasis; Hawley et al. (2018)) was significantly increased after the race, as well as the levels of fatty acids from lipoproteins and of certain amino acids, namely alanine, branched amino acids such as leucine, valine and *iso*-valerate, glutamate, glutamine and aromatic amino acids such as tyrosine and phenylalanine. Ketone bodies were slightly increased after the race (i.e., acetoacetate and acetate; **Supplementary Table S3**). In the case of biochemical profiles, all of the horses showed above-average concentrations for total bilirubin, creatine kinase, aspartate transaminase, and serum amyloid A after the race. The NEFA and BHB concentrations also showed similar patterns (**Supplementary Table S4**).

## Fecal Short Chain Fatty Acids Measurement, 16S rRNA Data and Microorganism Concentrations

The microbiota composition and derived-metabolites were obtained from Plancade et al. (2019), but it is important to note that the 16S rRNA raw sequences were re-analysed using the QIIME 2 plugin, which quantitatively improved results over QIIME 1 by enhancing the pre-processing of sequenced reads, the taxonomy assignment, the phylogenetic insertion and the generation of amplicon sequence variants (ASVs) (Callahan et al., 2017). A total of 519,866 high-quality sequence reads were obtained (mean per subject:  $21,131 \pm 15,625$ , range: 6,036–57,389). Reads were clustered into 3,385 chimera- and



singleton-filtered ASVs at 99% sequence similarity (**Supplementary Data Sheet 1**). The ASV taxonomical assignments and ASV counts for each individual are presented in the **Supplementary Table S6**). The intestinal microbial community found in the total set of 20 individuals as a whole was made up of a core of 23 genera, the core being defined as the genera shared by 99% of all sampling events with a minimum 0.1% mean relative abundance. Overall, 61% of the core genera belonged to the Firmicutes phylum, mainly to the *Lachnospiraceae* and *Ruminococcaceae* families (**Supplementary Figure S4A**). A total of 100 unique genera was identified in the microbiota (**Supplementary Table S7**). The majority of these genera (80%) fell among the 20 most abundant (**Supplementary Figure S4B**) and accounted for more than 75% of the sequences in the data (**Supplementary Table S7**). To deeper understand how the gut microbiota functioned, SCFAs, pH measurements and the loads of anaerobic fungi,

protozoa and total bacteria in feces were also investigated. The main products of microbial fermentation were acetate, propionate and butyrate, with ratios ranging from 60:32:8 to 76:19:5. Small amounts of branched chain fatty acids (*iso*-butyrate, valerate and *iso*-valerate) were also detected (**Supplementary Table S5**). Although bacteria represented the major portion of the fecal microbiota in our horses, the relative concentrations of anaerobic fungi and ciliate protozoa were 0.82 and 0.76, respectively (**Supplementary Table S8**).

### Integration of Transcriptome, <sup>1</sup>H NMR Metabolites, Biochemical Parameters, Fecal Microbiota, SCFA and Microorganism Concentrations

After the identification of the 801 mt-related genes that were regulated by endurance exercise, it remained to be determined

which genes were interconnected to the gut microbiota and responded to specific circulating molecules. To this aim, we applied four independent statistical methods using the mt-related genes as the response variable and the other data sets, namely  $^1\text{H}$  NMR metabolome, biochemical assay profiles, the filtered fecal microbiota (**Supplementary Table S9**), fecal SCFAs and the concentrations of bacteria, anaerobic fungi and protozoa as exploratory variables.

We first used global NMDS ordinations to visualize the structure of mt-related gene expression (ordination stress = 6%,  $k = 2$ , non-metric fit  $r^2 = 0.0996$ , linear fit  $r^2 = 0.988$ ) and we then fitted all sets of explanatory variables to the ordination to find the most influential variables (**Figure 3A**). Bacteria such as *Oribacterium*, *Rikenellaceae* RC9, *Ruminococcaceae* NK4A214, unclassified rumen bacterium and *Clostridium sensu stricto* showed the strongest correlation to all ordinations, together with some microbial-derived metabolites (i.e., dimethyl sulfone, formate, valerate and iso-valerate) and the plasmatic NEFA (adjusted  $p < 0.05$ ; **Figures 3A,B**).

To control for spatial variance and to identify the minimal combination of non-redundant covariates that would best fit with the mt-related gene profiles, a more rigorous multivariate distance-based redundancy analysis was used on constrained NMDS ordinations. In agreement with the aforementioned results, the expression of mt-related genes responded most strongly to bacteria such as *Treponema* ( $r_{\text{adj}}^2 = 0.48$ ,  $p = 0.007$ ), followed by *Butyrivibrio* ( $r_{\text{adj}}^2 = 0.45$ ,  $p = 0.002$ ), plasmatic NEFA ( $r_{\text{adj}}^2 = 0.38$ ,  $p = 0.003$ ), *Fibrobacter* ( $r_{\text{adj}}^2 = 0.31$ ,  $p = 0.0008$ ) and *Oribacterium* ( $r_{\text{adj}}^2 = 0.24$ ,  $p = 0.003$ ; **Figure 3C**). The  $^1\text{H}$  NMR metabolites and fecal SCFAs and the concentration of fecal microorganism did not directly contribute to the variation of mt-related genes. Therefore, they were not selected by the dbRDA model.

These findings were further confirmed by an RDA forward-selection model based on the Akaike information criterion. Specifically, *Oribacterium* ( $F = 7.26$ ,  $p < 0.005$ ), *Fibrobacter* ( $F = 2.96$ ,  $p < 0.005$ ), *Butyrivibrio* ( $F = 2.91$ ,  $p < 0.005$ ), *Aghostobacter* ( $F = 2.12$ ,  $p < 0.005$ ), *Treponema* ( $F = 7.15$ ,  $p < 0.01$ ), unclassified rumen bacterium ( $F = 1.96$ ,  $p < 0.01$ ), and the concentration of plasmatic NEFA ( $F = 2.91$ ,  $p < 0.005$ ) explained most of the variance observed in mt-related genes (**Figure 3D**). However, the  $^1\text{H}$  NMR metabolites, the fecal SCFAs and concentration of microorganisms were not found to contribute significantly to the variability in mt-related gene expression. The first constrained axis (RDA1) explained 44% of the variance in mt-related gene expression, and the second (RDA2) explained 9.4% of the variance; on the other hand, the two first unconstrained axes (PC1 and PC2) represented less than 8% of the total variance, i.e., much less than that explained by the explanatory variables together.

To uncover other potential underlying mechanisms of mitochondrial regulation, we then sought to examine the relationships existing among all aforementioned data sets by adding a further categorical variable, namely the racing performance of horses. To perform this task, we used the DIABLO framework from mixOmics (Singh et al., 2019). While the mt-related genes showed high levels of covariation

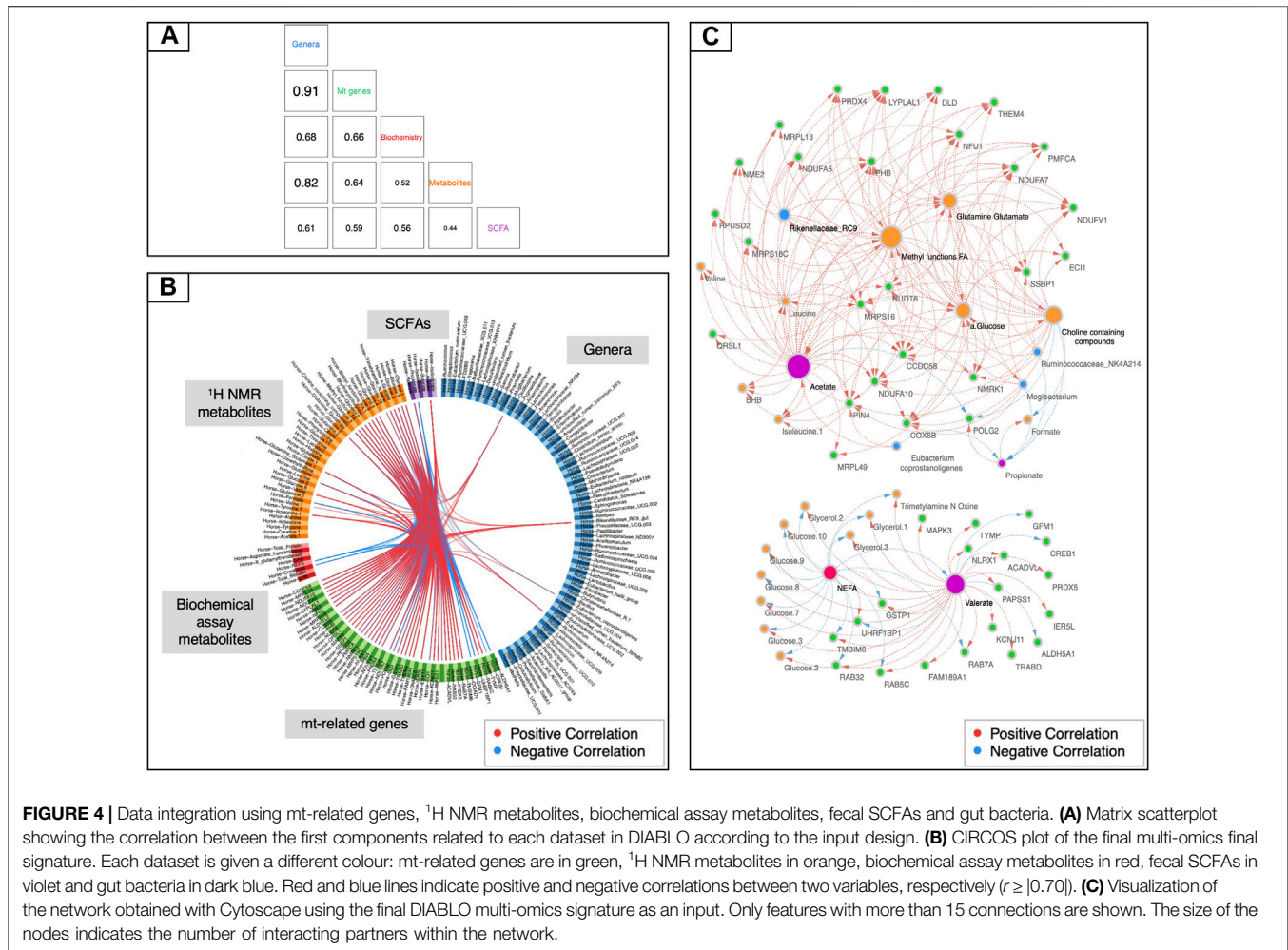
with the fecal microbiota ( $r^2 > 0.91$ , **Figure 4A**), it was not possible to identify a tight relationship with the other data sets. A more fine-grained view of this biological system was then obtained by focusing on pairwise correlations between variables. The first component of the DIABLO analysis highlighted a significant link between a subset of 45 mt-related genes (all encoded by the nuclear genome) and four gut taxa (i.e., the genus *Mogibacterium*, the species *Eubacterium coprostanoligenes* and the groups *Rikenellaceae* RC9 and *Ruminococcaceae* NK4A214). A link to blood metabolites related to energy supply (i.e., methyl groups of FAs and choline-containing compounds) and metabolites related to the TCA cycle such as glutamine, glutamate, and  $\alpha$ -glucose was also unveiled (**Figures 4B,C**). Moreover, mt-related genes co-occurred with pronounced variations of microbiota derived metabolites, including fecal acetate and valerate, and plasmatic concentrations of TMAO, dimethyl sulfone, and formate. The subset of 45 mt-related genes was functionally enriched in pathways related to fatty acid  $\beta$ -oxidation, mitochondrial apoptosis and biogenesis, respiratory electron transport and signaling and innate immune system response (**Supplementary Table S13**).

Finally, a rCCA analysis was also carried out to study in a more targeted way the relationships between mt-related genes and gut microbiota. In this case, the most relevant associations were represented by 90 genes and 9 bacterial genera (**Supplementary Table S14**). Overall, this method largely validated the associations already detected with the other aforementioned approaches. First, all of the four genera highlighted by DIABLO were confirmed, as well as the genus *Fibrobacter*, which had already been detected using NMDS and the db-RDA method. Second, three more taxa found with rCCA appeared to be functionally related to other previously identified microorganisms, thus providing indirect support to those findings. This was the case for *Ruminococcaceae* UCG-002, *Pseudobutyrvibrio* and the *Eubacterium hallii* group.

## DISCUSSION

In this work, we present an integrative study that combines the whole blood transcriptomic with untargeted serum metabolome data, blood biochemical assay profiles and gut metagenome in 20 equine-athletes. We assumed the adaptive response to extreme endurance exercise was essentially explained through the gut-mitochondria axis. Indeed, the different and complementary statistical approaches that we used confirmed this hypothesis, highlighting that the two main omic layers at play were the mt-related genes and the gut microbiota composition ( $r^2 > 0.91$ ).

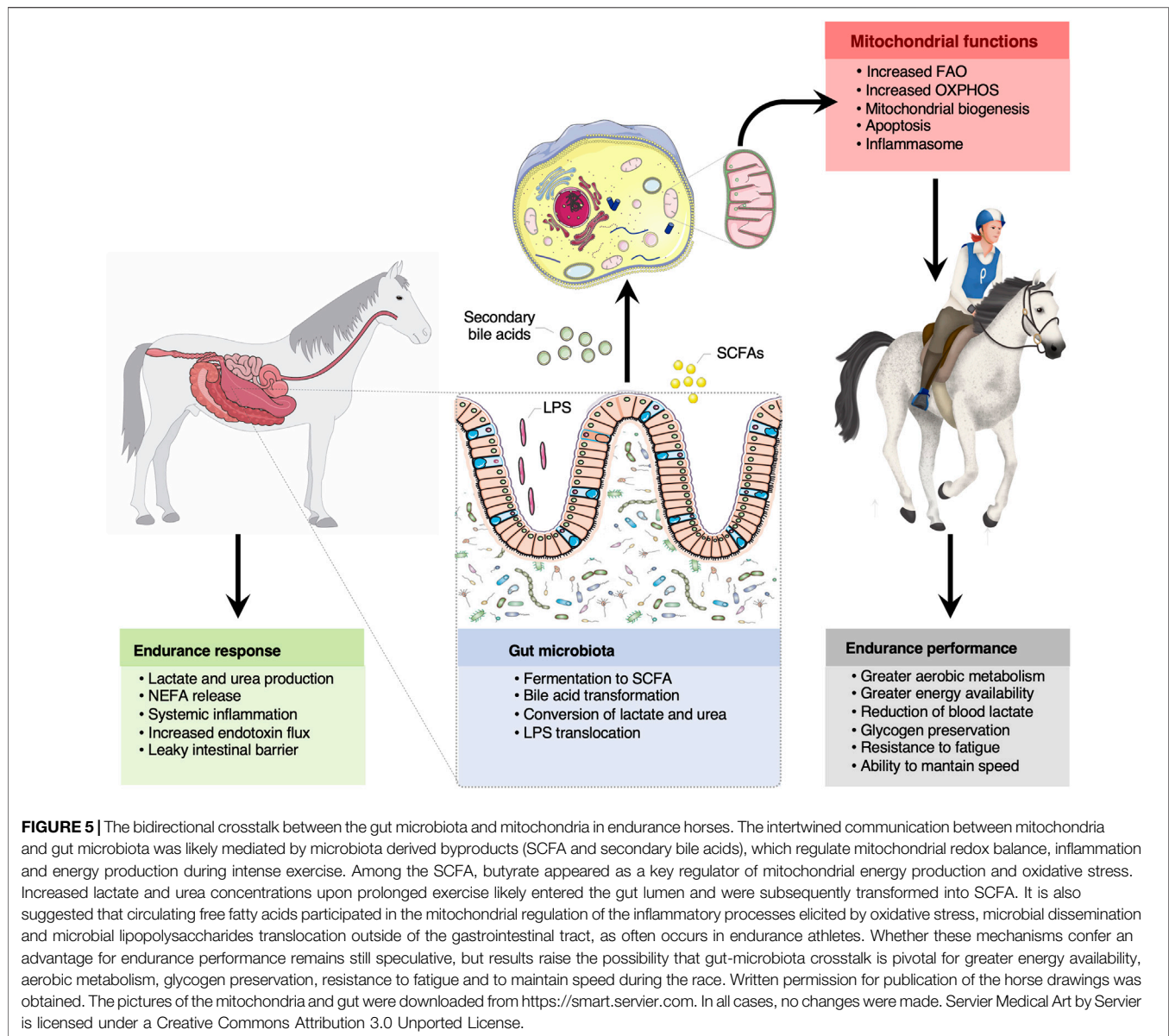
Whole blood transcriptome underlines in a clear manner the global response to exercise in equines (Barrey et al., 2006; Capomaccio et al., 2013; Mach et al., 2016; Mach et al., 2017b; Ropka-Molik et al., 2017), including the inflammatory response of the muscle associated with sarcolemma permeability and rhabdomyolysis (displayed in our study by the high levels of plasma creatine kinase and aspartate aminotransferase after the race). Yet, it remains to be explored whether the whole blood



transcriptome reflects the physiological events occurring at the mitochondrial level, notably in the tissues that are highly solicited under endurance, like for instance, the skeletal muscles, the heart and the liver (Gunn, 1987). In addition, whether the transcription of nuclearly-encoded and mitochondrially-encoded genes are regulated in a coordinated way is much less well understood. We therefore compared the transcriptome profile in equine-athletes before and after exercise and we used a meta-analytical approach that allowed us to identify 801 differentially expressed genes that were putatively linked to mitochondria. These genes fit neatly into the well-characterized context of adaptive mitochondrial regulation to endurance and mostly belonged to molecular pathways such as mitochondrial biogenesis, energy metabolism through OXPHOS and FAO, resistance to oxidative stress, mitophagy and inflammation regulation.

We then specifically focused on the mechanisms underlying the biological links between the aforementioned mt-related genes and the gut microbiota to untangle the gut-mitochondria crosstalk. The interdependence of mitochondria and gut microbiota is underscored by several lines of evidences (Mottawea et al., 2016; Gruber and Kennedy, 2017; Han et al.,

2017; Qi and Han, 2018; Saint-Georges-Chaumet and Edeas, 2018; Yardeni et al., 2019; Ruiz et al., 2020; Zhang et al., 2020), although the range and extent of this interplay are largely unknown. For example, Mottawea et al. (2016) showed that butyrate-producing bacteria and mitochondrial proteins were positively correlated, suggesting a signaling role for butyrate in mitochondrial gene expression. In support of this observation, our results revealed that several functionally redundant butyrate-producing bacterial families were associated with the mt-related genes, namely *Lachnospiraceae* (*Oribacterium*, *Butyrivibrio*, *Agathobacter* and *Eubacterium* spp.), *Ruminococcaceae*, *Spirochaetaceae* (*Treponema* spp.) (Vital et al., 2015; Gharechahi et al., 2020; Vacca et al., 2020) and *Rikenellaceae* (Vital et al., 2015). The bioavailability of butyrate is obviously related to endurance performance because of the role played by this molecule in energy metabolism (Mollica et al., 2017). Beyond the scope of its energy producing capacity, butyrate is also known to induce the expression of *PPAR $\gamma$*  gene (Gao et al., 2009) and downstream targets in different cells. Our blood transcriptomic analysis indicated an upregulation of *PPAR $\gamma$*  following exercise, raising the possibility that the enrichment in butyrate-producing bacteria



increased the expression of this transcription factor and downstream signaling, leading to fatty acid shuttling into and oxidation by the mitochondria. Notably, the redox imbalance during strenuous exercise might be also attenuated by butyrate (Dobashi et al., 2011; Mottawea et al., 2016). The potential of butyrate to improve exercise capacity has been further posited by Gao et al. (2009) and Henagan et al. (2015), who observed that the supplementation of this molecule improved the oxidative skeletal muscle phenotype, its mitochondrial content and its proportion of type I fibers. Concomitantly, it is possible that serum lactate, which appeared to be significantly increased in our horses upon prolonged exercise, entered the gut lumen, where it was subsequently transformed into butyrate by *Eubacterium hallii* (Duncan et al., 2004; Scheiman et al., 2019). Indeed, *Eubacterium hallii* was associated to a subset of mt-related genes according to our rCCA analysis. *Eubacterium*-derived butyrate could then be

absorbed into the portal vein and serve as an energy source to the different organs. Another plausible mechanisms employed by microbiota to communicate with mitochondria involved valerate, a branched SCFA formed from protein and amino acid degradation (Fernandes et al., 2014). It may be debated whether the urea produced by the host during endurance could be hydrolyzed by commensal gut microbiota resulting in valerate.

Beyond SCFAs, the secondary bile acids could also play an important role in gut-microbiota crosstalk. The genera *Eubacterium* and *Clostridium*, which contributed significantly to our biological system, have the capacity to degrade 5–10% of the primary bile acids forming secondary bile acids (Gérard, 2013). Secondary bile acids might interact with the mitochondria via the activation *FXR-CREB* axis. *CREB*, which was significantly increased in our horses after the endurance race, is a sensor of

energy charge and other stress signals, and is a regulator of metabolism that activates autophagy and lipid catabolic functions (Seok et al., 2014). Therefore, a picture emerges that, under conditions that foster an increased colonization by these microorganisms, the production of butyrate, valerate and secondary bile acids in the intestine is likely increased, with potential effects on the mitochondria functionality and endurance performance (Figure 5). Yet, evidence of causality of microbiome-derived metabolites on the gut-mitochondria crosstalk remains elusive.

The coordination between mitochondria and gut microbiota was presumably regulated by the circulating free fatty acids, the so-called NEFAs. It is becoming clearer that NEFAs not only serves as energy source in the working muscles but act as extracellular signaling molecules that modulate the production of chemokines and cytokines, and the synthesis of pro-inflammatory lipid-derived species (Rodríguez-Carrio et al., 2017). Thus, a provocative extension of our work suggests that increased release of NEFAs participated in the mitochondrial regulation of the inflammatory processes elicited by oxidative stress, microbial dissemination and microbial lipopolysaccharides translocation outside of the gastrointestinal tract, commonly observed in endurance athletes (Fielding and Dechant, 2012). Increased release of free fatty acids may dampen the inflammatory response and prevent or mitigate the negative effects of redox imbalance. Supporting this notion, the mitochondrial sirtuin (*SIRT5*) and *SIRT1*, which have an anti-inflammation function (Wang et al., 2017), were found to be increased during endurance.

The observations presented herein indicate that the horse could be considered as an interesting *in vivo* model for research in the field of human exercise given its large body size, the aptitude for endurance exercise (Votion et al., 2012; van der Kolk et al., 2020), that is, high baseline maximal oxygen uptake ( $VO_{2max}$ ;  $\sim 120 \text{ ml min}^{-1} \text{ kg}^{-1}$ ) and the ability to sustain work at a high percentage of  $VO_{2max}$  without either the accumulation of exponential levels of blood lactate or skeletal muscle fatigue and the exercise economy (Cottin et al., 2010; Goachet and Julliand, 2015; van der Kolk et al., 2020). The level of exercise performed by a horse during an endurance competition is similar to that of a human marathon runner (Capomaccio et al., 2013; Mach et al., 2016) or ultramarathon runner (Scott et al., 2009). Nevertheless, despite the usefulness of this model, the differences between the microbiota of horses and humans are relevant. In contrast to what happens in humans, in horses the cecum is large relative to the total gastrointestinal tract and it is an important site for the fermentation of plant materials. Our study presents other limitations. Although the omic approaches used here are considered robust and generate high quality data, they still present several limitations. For instance, 16S rRNA sequencing measures the relative abundance of bacterial genera contained in it, but it does not give any information about its actual functionality, which should be therefore evaluated using other methods, such as for instance metatranscriptomics. Moreover, in our case,  $^1\text{H}$  NMR has been able to detect only metabolites at high concentrations, like in the cases of amino acids, lipids, choline and N-acetylglucosamine. In this regard, the

combination of  $^1\text{H}$  NMR and mass spectrometry should result in better coverage of metabolites derived from bacteria, metabolites that are produced by the host and then modified by bacteria and metabolites that are *de novo* synthesized by bacteria. Lastly, it still remains to be determined how the individual components of blood, including plasma, platelets, erythrocytes, nucleated blood cells and exosomes reflect the transcriptomic profiles in horses. Upon endurance, contracting muscles release proteins and metabolites that have endocrine-like properties (Hawley, 2020), but they might also release long non-coding RNAs, myo-miRs and circulating cell-free respiratory competent mitochondria (Al Amir Dache et al., 2020; Song et al., 2020) that might participate in the aforementioned crosstalk by using the microbiota-derived metabolites.

Taken together, the present study offers extensive novel insight into the mitochondria-gut microbiota axis and opens the way for mechanistic studies that will lead to a better understanding of the orchestrated molecular pathways that underpin endurance adaptations and contribute to the holobiont biology. This is the first description of how metabolites derived from commensal gut microbiota (SCFAs and secondary bile acids), or produced by the host and biochemically modified by gut bacteria (lactate and urea), might influence the genes related to the mitochondria and involved in energy production, redox balance and inflammatory cascades, making them a potential therapeutic target for the endurance. The activation of *PPAR $\gamma$*  and the *FRX-CREB* axis are likely key mechanisms through which SCFAs and bile acids coordinately engage multiple converging pathways to regulate mitochondrial functions, including fatty acid uptake and oxidation to forestall hypoglycemia and ensure longer running time.

Our results also suggest that free fatty acids may not only serve as an important fuel for skeletal muscle during endurance, but may also regulate mitochondrial inflammatory responses through a plethora of mechanisms, the principal one likely being the modulation of the intestinal barrier-ROS production and lipopolysaccharide translocation. Further research focusing on the role that gut microbiota plays on the mitochondrial function across a wide range of tissues and cell types may be highly informative to improve the athlete's energy metabolism, redox status and inflammatory response.

## DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. This data can be found here: Microarray expression data: Gene Expression Omnibus (GEO) repository under the accession number GSE163767 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE163767>) Metabolomic data: NIH Common Fund's Data Repository and Coordinating Center UrqK1489; (<http://dev.metabolomicsworkbench.org:2222/data/DRCCMetadata.php?Mode=Study&StudyID=ST000945>) Gut metagenome 16S rRNA targeted locus data: DDBJ/EMBL/GenBank under the accession KBTQ0000000, version KBTQ0000000.1; (locus KBTQ01000000). The corresponding BioProject is PRJNA438436, and the accession numbers of the BioSamples included in it were SAMN08715709 to SAMN08715760.

## ETHICS STATEMENT

The study protocol was reviewed and approved by the local Animal Care and Use Committee (ComEth EnvA-Upec-ANSES, reference: 11-0041, dated July 12th, 2011) for horse study. All the protocols were conducted in accordance with EEC regulation (no 2010/63/UE) governing the care and use of laboratory animals, which has been effective in France since the 1st of January 2013. In all cases, the owners and riders provided their informed consent prior to the start of study procedures with the animals. Written informed consent was obtained from the owners for the participation of their animals in this study.

## AUTHOR CONTRIBUTIONS

NM and EB conceived of the presented idea. NM designed the experiment, performed the PCR and the RT-qPCR experiments, carried out most of the bioinformatics analysis and the integrative biology approaches. NM wrote the main manuscript text and prepared the figures with support from MM and EB. MM carried out the transcriptome analyses, the rCCA integrative analysis and prepared most of the tables. MM and EB created the consensus list of genes related to mitochondria. AR verified the statistical methods and CR verified the blood biochemical and the performance data. JL performed all the laboratory steps related to microarray analyses. LL performed the metabolomic experiment and analyzed the metabolite peaks. CR and EB were in charge of the organization and sampling management during the race. All authors reviewed the manuscript and approved the final version.

## REFERENCES

- Al Amir Dache, Z., Otandault, A., Tanos, R., Pastor, B., Meddeb, R., Sanchez, C., et al. (2020). Blood contains circulating cell-free respiratory competent mitochondria. *FASEB J.* 34, 3616–3630. doi:10.1096/fj.201901917RR
- Bajpai, P., Darra, A., and Agrawal, A. (2018). Microbe-mitochondrion crosstalk and health: an emerging paradigm. *Mitochondrion* 39, 20–25. doi:10.1016/j.mito.2017.08.008
- Barrey, E., Mucher, E., Robert, C., Amiot, F., and Gidrol, X. (2006). Gene expression profiling in blood cells of endurance horses completing competition or disqualified due to metabolic disorder. *Equine Vet. J.* 38, 43–49. doi:10.1111/j.2042-3306.2006.tb05511.x
- Barton, W., Penney, N. C., Cronin, O., Garcia-Perez, I., Molloy, M. G., Holmes, E., et al. (2017). The microbiome of professional athletes differs from that of more sedentary subjects in composition and particularly at the functional metabolic level. *Gut.* 31, 267. doi:10.1136/gutjnl-2016-313627
- Bindea, G., Mlecnik, B., Hackl, H., Charoentong, P., Tosolini, M., Kirilovsky, A., et al. (2009). ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics* 25, 1091–1093. doi:10.1093/bioinformatics/btp101
- Blanchet, F. G., Legendre, P., and Borcard, D. (2008). Forward selection of explanatory variables. *Ecology* 89, 2623–2632. doi:10.1890/07-0986.1
- Bokulich, N. A., Kaehler, B. D., Rideout, J. R., Dillon, M., Bolyen, E., Knight, R., et al. (2018). Optimizing taxonomic classification of marker-gene amplicon

## FUNDING

This work was funded by grants from the *Fonds Eperon*, the *Institut Français du Cheval et de l'Équitation* (IFCE), the *Association du Cheval Arabe* (ACA) and the *Institute National de la Recherche pour l'Agriculture, l'alimentation et l'environnement* (INRAE).

## ACKNOWLEDGMENTS

We are grateful to Marine Beinat, Julie Rivière, Emmanuelle Rebours, Jordi Estellé, Caroline Morgenthaler and Arni Janssen for participating in the sample collection and organization during the project. We particularly thank Valentine Ballan and Fanny Blanc who helped to isolate the blood RNA. We also thank Catherine Philippe for the SCFA analysis in feces, Diane Esquerré who prepared the libraries and performed the MiSeq sequencing at the GeT-PlaGe genomics core facility (INRAE, Toulouse, <https://get.genotoul.fr/>) and Estel Blasi, who created the picture of horses. Lastly, we are grateful to the INRAE MIGALE bioinformatics facility (MIGALE, INRAE, 2020, Migale bioinformatics facility, doi: 10.15454/1.5572390655343293E12) for providing computing and storage resources and to all the horse owners, riders and endurance race organizers who participated in the study.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmolb.2021.656204/full#supplementary-material>.

sequences with QIIME 2's q2-feature-classifier plugin. *Microbiome* 6, 90. doi:10.1186/s40168-018-0470-z

- Callahan, B. J., McMurdie, P. J., and Holmes, S. P. (2017). Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *ISME J.* 11, 2639–2643. doi:10.1038/ismej.2017.119
- Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., and Holmes, S. P. (2016). DADA2: high-resolution sample inference from Illumina amplicon data. *Nat. Methods* 13, 581. doi:10.1038/nmeth.3869
- Calvo, S. E., Clauser, K. R., and Mootha, V. K. (2016). MitoCarta2.0: an updated inventory of mammalian mitochondrial proteins. *Nucl. Acids Res.* 44, D1251–D1257. doi:10.1093/nar/gkv1003
- Capomaccio, S., Vitulo, N., Verini-Supplizi, A., Barcaccia, G., Albiero, A., D'Angelo, M., et al. (2013). RNA sequencing of the exercise transcriptome in equine athletes. *PLoS One* 8, e83504. doi:10.1371/journal.pone.0083504
- Cardoso, S. M., and Empadinhas, N. (2018). The microbiome-mitochondria dance in prodromal Parkinson's disease. *Front. Physiol.* 9, 471. doi:10.3389/fphys.2018.00471
- Chinnery, P. F., and Hudson, G. (2013). Mitochondrial genetics. *Br. Med. Bull.* 106, 135–139. doi:10.1093/bmb/ldt017
- Clark, A., and Mach, N. (2017). The crosstalk between the gut microbiota and mitochondria during exercise. *Front. Physiol.* 8, 319. doi:10.3389/fphys.2017.00319
- Clark, A., Sallé, G., Ballan, V., Reigner, F., Meynadier, A., Cortet, J., et al. (2018). Strongyle infection and gut microbiota: profiling of resistant and susceptible horses over a grazing season. *Front. Physiol.* 9, 1–18. doi:10.3389/fphys.2018.00272



- Clarke, K. R., and Ainsworth, M. (1993). A method of linking multivariate community structure to environmental variables. *Mar. Ecol. Prog. Ser.* 92, 205–219. doi:10.3354/meps092205
- Costa, M. C., and Weese, J. S. (2018). Understanding the intestinal microbiome in health and disease. *Vet. Clin. North Am. Equine Pract.* 34, 1–12.
- Cottin, F., Metayer, N., Goachet, A. G., Julliand, V., Slawinski, J., Billat, V., et al. (2010). Oxygen consumption and gait variables of Arabian endurance horses measured during a field exercise test. *Equine Vet. J.* 42, 1–5. doi:10.1111/j.2042-3306.2010.00184.x
- Crisp, A., Boschetti, C., Perry, M., Tunnacliffe, A., and Micklem, G. (2015). Expression of multiple horizontally acquired genes is a hallmark of both vertebrate and invertebrate genomes. *Genome Biol.* 16, 50. doi:10.1186/s13059-015-0607-3
- Dieterle, F., Ross, A., Schlotterbeck, G., and Senn, H. (2006). Probabilistic quotient normalization as robust method to account for dilution of complex biological mixtures. Application in <sup>1</sup>H NMR metabolomics. *Anal. Chem.* 78, 4281–4290. doi:10.1021/ac051632c
- Dixon, P. (2003). VEGAN, a package of R functions for community ecology. *J. Veg. Sci.* 14, 927. doi:10.1658/1100-9233(2003)014[0927:VAPORF]2.0.CO;2
- Dobashi, Y., Miyakawa, Y., Yamamoto, I., and Amao, H. (2011). Effects of intestinal microflora on superoxide dismutase activity in the mouse cecum. *Exp. Anim.* 60, 133–139. doi:10.1538/expanim.60.133
- Duncan, S. H., Louis, P., and Flint, H. J. (2004). Lactate-utilizing bacteria, isolated from human feces, that produce butyrate as a major fermentation product. *Appl. Environ. Microbiol.* 70, 5810–5817. doi:10.1128/AEM.70.10.5810-5817.2004
- Esser, C., Ahmadinejad, N., Wiegand, C., Rotte, C., Sebastiani, F., Gelius-Dietrich, G., et al. (2004). A genome phylogeny for mitochondria among  $\alpha$ -proteobacteria and a predominantly eubacterial ancestry of yeast nuclear genes. *Mol. Biol. Evol.* 21, 1643–1660. doi:10.1093/molbev/msh160
- Fernandes, J., Su, W., Rahat-Rozenbloom, S., Wolever, T. M. S., and Comelli, E. M. (2014). Adiposity, gut microbiota and faecal short chain fatty acids are linked in adult humans. *Nutr. Diabetis* 4, e121. doi:10.1038/ntud.2014.23
- Fielding, C. L., and Dechant, J. E. (2012). Colic in competing endurance horses presenting to referral centres: 36 cases. *Equine Vet. J.* 44, 472–475. doi:10.1111/j.2042-3306.2011.00462.x
- Franco-Obregón, A., and Gilbert, J. A. (2017). The microbiome-mitochondrion connection: common ancestries, common mechanisms, common goals. *mSystems* 2, e00018–17. doi:10.1128/mSystems.00018-17
- Gao, Z., Yin, J., Zhang, J., Ward, R. E., Martin, R. J., Lefevre, M., et al. (2009). Butyrate improves insulin sensitivity and increases energy expenditure in mice. *Diabetes* 58, 1509–1517. doi:10.2337/db08-1637
- Gérard, P. (2013). Metabolism of cholesterol and bile acids by the gut microbiota. *Pathogens* 3, 14–24. doi:10.3390/pathogens3010014
- Gharechahi, J., Farhad, M., Mohammad, V., and Han, J. (2020). Metagenomic analysis reveals a dynamic microbiome with diverse and adaptive functions to utilize high lignocellulosic forages in the cattle rumen. *ISME J.* doi:10.1038/s41396-020-00837-2
- Goachet, A. G., and Julliand, V. (2015). Implementation of field cardio-respiratory measurements to assess energy expenditure in Arabian endurance horses. *Animal* 9, 787–792. doi:10.1017/S1751731114003061
- Gruber, J., and Kennedy, B. K. (2017). Microbiome and longevity: gut microbes send signals to host mitochondria. *Cell* 169, 1168–1169. doi:10.1016/j.cell.2017.05.048
- Gunn, H. M. (1987). “Muscle, bone and fat proportions and muscle distribution of thoroughbreds and other horses,” in *International conference on equine exercise physiology*. Editors J. R. Gillespie and N. E. Robinson (Davis, CA: ICEEP Publications), 253–264.
- Han, B., Sivaramkrishnan, P., Lin, C. C. J., Neve, I. A. A., He, J., Tay, L. W. R., et al. (2017). Microbial genetic composition tunes host longevity. *Cell* 169, 1249–1262. doi:10.1016/j.cell.2017.05.036
- Hawley, J. A., Lundby, C., Cotter, J. D., and Burke, L. M. (2018). Maximizing cellular adaptation to endurance exercise in skeletal muscle. *Cell Metab* 27, 962–976. doi:10.1016/j.cmet.2018.04.014
- Hawley, J. A. (2020). Microbiota and muscle highway—two way traffic. *Nat. Rev. Endocrinol.* 16, 71–72. doi:10.1038/s41574-019-0291-6
- Henagan, T. M., Stefanska, B., Fang, Z., Navard, A. M., Ye, J., Lenard, N. R., et al. (2015). Sodium butyrate epigenetically modulates high-fat diet-induced skeletal muscle mitochondrial adaptation, obesity and insulin resistance through nucleosome positioning. *Br. J. Pharmacol.* 172, 2782–2798. doi:10.1111/bph.13058
- Hood, D. A., Ugucioni, G., Vainshtein, A., and D’souza, D. (2011). Mechanisms of exercise-induced mitochondrial biogenesis in skeletal muscle: implications for health and disease. *Compr. Physiol.* 1, 1119–1134. doi:10.1002/cphy.c100074
- Jackson, D. N., and Theiss, A. L. (2020). Gut bacteria signaling to mitochondria in intestinal inflammation and cancer. *Gut Microbes* 11, 285–304. doi:10.1080/19490976.2019.1592421
- Jang, H. J., Kim, D. M., Kim, K. B., Park, J. W., Choi, J. Y., Hyeog Oh, J., et al. (2017). Analysis of metabolomic patterns in thoroughbreds before and after exercise. *Asian-Australas. J. Anim. Sci.* 30, 1633–1642. doi:10.5713/ajas.17.0167
- Janouškovec, J., Tikhonenkov, D. V., Burki, F., Howe, A. T., Rohwer, F. L., Mylnikov, A. P., et al. (2017). A new lineage of eukaryotes illuminates early mitochondrial genome reduction. *Curr. Biol.* 27, 3717–3724. doi:10.1016/j.cub.2017.10.051
- Jones, R. M., and Neish, A. S. (2014). Redox signaling mediates symbiosis between the gut microbiota and the intestine. *Gut Microbes* 5, 250–253. doi:10.4161/gmic.27917
- Karlsson, F. H., Tremaroli, V., Nookaew, I., Bergström, G., Behre, C. J., Fagerberg, B., et al. (2013). Gut metagenome in European women with normal, impaired and diabetic glucose control. *Nature* 498, 99–103. doi:10.1038/nature12198
- Kauffmann, A., Gentleman, R., and Huber, W. (2009). arrayQualityMetrics - a bioconductor package for quality assessment of microarray data. *Bioinformatics* 25, 415–416. doi:10.1093/bioinformatics/btn647
- Kauter, A., Epping, L., Semmler, T., Antao, E.-M., Kannapin, D., Stoeckle, S. D., et al. (2019). The gut microbiome of horses: current research on equine enteric microbiota and future perspectives. *Anim. Microbiome* 1, 1–15. doi:10.1186/s42523-019-0013-3
- Keohane, D. M., Woods, T., O’Connor, P., Underwood, S., Cronin, O., Whiston, R., et al. (2019). Four men in a boat: ultra-endurance exercise alters the gut microbiome. *J. Sci. Med. Sport* 22, 1059–1064. doi:10.1016/j.jsams.2019.04.004
- Langfelder, P., and Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* 9, 559. doi:10.1186/1471-2105-9-559
- Le Moyec, L., Robert, C., Triba, M. N., Billat, V. L., Mata, X., Schibler, L., et al. (2014). Protein catabolism and high lipid metabolism associated with long-distance exercise are revealed by plasma NMR metabolomics in endurance horses. *PLoS One* 9, 1–10. doi:10.1371/journal.pone.0090730
- Le Moyec, L., Robert, C., Triba, M. N., Bouchemal, N., Mach, N., Rivière, J., et al. (2019). A first step toward unraveling the energy metabolism in endurance horses: comparison of plasma nuclear magnetic resonance metabolomic profiles before and after different endurance race distances. *Front. Mol. Biosci.* 6, 45. doi:10.3389/fmolb.2019.00045
- Lê, S., Josse, J., and Husson, F. (2008). FactoMineR: an R package for multivariate analysis. *J. Stat. Softw.* 25, 1–18. doi:10.1016/j.envint.2008.06.007
- Mach, N., Berri, M., Estellé, J., Levenez, F., Lemonnier, G., Denis, C., et al. (2015). Early-life establishment of the swine gut microbiome and impact on host phenotypes. *Environ. Microbiol. Rep.* 7, 554–569. doi:10.1111/1758-2229.12285
- Mach, N., Foury, A., Kittelmann, S., Reigner, F., Moroldo, M., Ballester, M., et al. (2017a). The effects of weaning methods on gut microbiota composition and horse physiology. *Front. Physiol.* 8, 535. doi:10.3389/fphys.2017.00535
- Mach, N., and Fuster-Botella, D. (2017). Endurance exercise and gut microbiota: a review. *J. Sport Heal. Sci.* 6, 179–197. doi:10.1016/j.jshe.2016.05.001
- Mach, N., Plancade, S., Pacholewska, A., Lecardonnell, J., Rivière, J., Moroldo, M., et al. (2016). Integrated mRNA and miRNA expression profiling in blood reveals candidate biomarkers associated with endurance exercise in the horse. *Sci. Rep.* 6, 22932. doi:10.1038/srep22932
- Mach, N., Ramayo-Caldas, Y., Clark, A., Moroldo, M., Robert, C., Barrey, E., et al. (2017b). Understanding the response to endurance exercise using a systems biology approach: combining blood metabolomics, transcriptomics and miRNomics in horses. *BMC Genomics* 18, 187. doi:10.1186/s12864-017-3571-3
- Mach, N., Ruet, A., Clark, A., Bars-Cortina, D., Ramayo-Caldas, Y., Crisci, E., et al. (2020). Priming for welfare: gut microbiota is associated with equitation conditions and behavior in horse athletes. *Sci. Rep.* 10, 8311.
- Maier, U. G., Zauner, S., Woehle, C., Bolte, K., Hempel, F., Allen, J. F., et al. (2013). Massively convergent evolution for ribosomal protein gene content in plastid and mitochondrial genomes. *Genome Biol. Evol.* 12, 2318–2329. doi:10.1093/gbe/evt181
- Massacci, F. R., Clark, A., Ruet, A., Lansade, L., Costa, M., and Mach, N. (2019). Inter-breed diversity and temporal dynamics of the faecal microbiota in healthy horses. *J. Anim. Breed. Genet.* 00, 1–18. doi:10.1111/jbg.12441
- McMurdie, P. J., and Holmes, S. (2013). Phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One* 8, e61217.

- Mcmurdie, P. J., and Holmes, S. (2012). Phyloseq: a bioconductor package for handling and analysis of high-throughput phylogenetic sequence data. *Pac. Symp. Biocomput* 10, 235–246.
- Mohr, S., and Liew, C. C. (2007). The peripheral-blood transcriptome: new insights into disease and risk assessment. *Trends Mol. Med.* 13, 422–432. doi:10.1016/j.molmed.2007.08.003
- Mollica, M. P., Raso, G. M., Cavaliere, G., Trinchese, G., De Filippo, C., Aceto, S., et al. (2017). Butyrate regulates liver mitochondrial function, efficiency, and dynamics in insulin-resistant obese mice. *Diabetes* 66, 1405–1418. doi:10.2337/db16-0924
- Mottawea, W., Chiang, C. K., Mühlbauer, M., Starr, A. E., Butcher, J., Abujamel, T., et al. (2016). Altered intestinal microbiota-host mitochondria crosstalk in new onset Crohn's disease. *Nat. Commun.* 7, 13419. doi:10.1038/ncomms13419
- Nay, K., Jollet, M., Goustard, B., Baati, N., Vernus, B., Pontones, M., et al. (2019). Gut bacteria are critical for optimal muscle function: a potential link with glucose homeostasis. *Am. J. Physiol. Endocrinol. Metab.* 317, E158–171. doi:10.1152/ajpendo.00521.2018
- Okamoto, T., Morino, K., Ugi, S., Nakagawa, F., Lemecha, M., Ida, S., et al. (2019). Microbiome potentiates endurance exercise through intestinal acetate production. *Am. J. Physiol. Endocrinol. Metab.* 316, E956–E966. doi:10.1152/ajpendo.00510.2018
- Oksanen, J., Blanchet, F. G., Kindt, R., Minchin, P. R., O'Hara, R. B., Simpson, G. L., et al. (2013). *Package 'vegan': a community ecology package*, 260. doi:10.1111/j.1654-1103.2003.tb02228.x
- Pagliarini, D. J., Calvo, S. E., Chang, B., Sheth, S. A., Vafai, S. B., Ong, S. E., et al. (2008). A mitochondrial protein compendium elucidates complex I disease biology. *Cell* 134, 112–123. doi:10.1016/j.cell.2008.06.016
- Pfanner, N., Warscheid, B., and Wiedemann, N. (2019). Mitochondrial proteins: from biogenesis to functional networks. *Nat. Rev. Mol. Cell Biol.* 20, 267–284. doi:10.1038/s41580-018-0092-0
- Plancade, S., Clark, A., Philippe, C., Helbling, J.-C., Moisan, M.-P., Esquerré, D., et al. (2019). Unraveling the effects of the gut microbiota composition and function on horse endurance physiology. *Sci. Rep.* 9, 9620.
- Qi, B., and Han, M. (2018). Microbial siderophore enterobactin promotes mitochondrial iron uptake and development of the host via interaction with ATP synthase. *Cell* 175, 571–582. doi:10.1016/j.cell.2018.07.032
- Rapoport, B. I. (2010). Metabolic factors limiting performance in marathon runners. *PLoS Comput. Biol.* 6, e1000960. doi:10.1371/journal.pcbi.1000960
- Richter-Dennerlein, R., Oeljeklaus, S., Lorenzi, I., Ronsör, C., Bareth, B., Schendzielorz, A. B., et al. (2016). Mitochondrial protein synthesis adapts to influx of nuclear-encoded protein. *Cell* 167, 471–483. doi:10.1016/j.cell.2016.09.003
- Rodríguez-Carrión, J., Salazar, N., Margolles, A., González, S., Gueimonde, M., de los Reyes-Gavilán, C. G., et al. (2017). Free fatty acids profiles are related to gut microbiota signatures and short-chain fatty acids. *Front. Immunol.* 8, 823. doi:10.3389/fimmu.2017.00823
- Rohart, F., Gautier, B., Singh, A., and Lê Cao, K. A. (2017). mixOmics: an R package for omics feature selection and multiple data integration. *PLoS Comput. Biol.* 13, e1005752. doi:10.1371/journal.pcbi.1005752
- Ropka-Molik, K., Stefaniuk-Szmukier, M., Zukowski, K., Piórkowska, K., Gurgul, A., and Bugno-Poniewierska, M. (2017). Transcriptome profiling of Arabian horse blood during training regimens. *BMC Genet.* 18, 31. doi:10.1186/s12863-017-0499-1
- Ruiz, E., Penrose, H. M., Heller, S., Nakhoul, H., Baddoo, M., Flemington, E. F., et al. (2020). Bacterial TLR4 and NOD2 signaling linked to reduced mitochondrial energy function in active inflammatory bowel disease. *Gut Microbes* 11, 350–363. doi:10.1080/19490976.2019.1611152
- Saint-Georges-Chaumet, Y., Attaf, D., Pelletier, E., and Edeas, M. (2015). Targeting microbiota-mitochondria inter-talk: microbiota control mitochondria metabolism. *Cell. Mol. Biol.* 61, 121–124. doi:10.14715/cmb/2015.61.4.20
- Saint-Georges-Chaumet, Y., and Edeas, M. (2018). Microbiota-mitochondria inter-talk: consequence for microbiota-host interaction. *Pathog. Dis.* 74, 1–5. doi:10.1093/femspd/ftv096
- Scheiman, J., Luber, J. M., Chavkin, T. A., MacDonald, T., Tung, A., Pham, L. D., et al. (2019). Meta-omics analysis of elite athletes identifies a performance-enhancing microbe that functions via lactate metabolism. *Nat. Med.* 25, 1104–1109. doi:10.1038/s41591-019-0485-4
- Scott, J. M., Esch, B. T. A., Shave, R., Warburton, D. E. R., Gaze, D., and George, K. (2009). Cardiovascular consequences of completing a 160-km ultramarathon. *Med. Sci. Sports Exerc.* 41, 25–33. doi:10.1249/MSS.0b013e31818313ff
- Seok, S., Fu, T., Choi, S. E., Li, Y., Zhu, R., Kumar, S., et al. (2014). Transcriptional regulation of autophagy by an FXR-CREB axis. *Nature* 516, 108–111. doi:10.1038/nature13949
- Singh, A., Shannon, C. P., Gautier, B., Rohart, F., Vacher, M., Tebbutt, S. J., et al. (2019). DIABLO: an integrative approach for identifying key molecular drivers from multi-omics assays. *Bioinformatics* 35, 3055–3062. doi:10.1093/bioinformatics/bty1054
- Smith, A. C., and Robinson, A. J. (2019). Mitominer v4.0: an updated database of mitochondrial localization evidence, phenotypes and diseases. *Nucleic Acids Res.* 47, D1225–D1228. doi:10.1093/nar/gky1072
- Smyth, G. K. (2004). Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat. Appl. Genet. Mol. Biol.* 3. doi:10.2202/1544-6115.1027
- Smyth, G. K., Michaud, J., and Scott, H. S. (2005). Use of within-array replicate spots for assessing differential expression in microarray experiments. *Bioinformatics* 21, 2067–2075. doi:10.1093/bioinformatics/bti270
- Song, X., Hu, W., Yu, H., Wang, H., Zhao, Y., Korngold, R., et al. (2020). Existence of circulating mitochondria in human and animal peripheral blood. *Int. J. Mol. Sci.* 21, 2122. doi:10.3390/ijms21062122
- Vacca, M., Celano, G., Calabrese, F. M., Portincasa, P., Gobetti, M., and De Angelis, M. (2020). The controversial role of human gut lachnospiraceae. *Microorganisms* 8, 1–25. doi:10.3390/microorganisms8040573
- van der Kolk, J. H., Thomas, S., Mach, N., Ramseyer, A., Burger, D., Gerber, V., et al. (2020). Serum acylcarnitine profile in endurance horses with and without metabolic dysfunction. *Vet. J.* 255, 105419. doi:10.1016/j.tvjl.2019.105419
- Veza, T., Abad-Jiménez, Z., Martí-Cabrera, M., Rocha, M., and Víctor, V. M. (2020). Microbiota-mitochondria inter-talk: a potential therapeutic strategy in obesity and type 2 diabetes. *Antioxidants* 9, 1–21. doi:10.3390/antiox9090848
- Vital, M., Gao, J., Rizzo, M., Harrison, T., and Tiedje, J. M. (2015). Diet is a major factor governing the fecal butyrate-producing community structure across Mammalia, Aves and Reptilia. *ISME J.* 9, 832–843. doi:10.1038/ismej.2014.179
- Votien, D. M., Gnaiger, E., Lemieux, H., Mouithys-Mickalad, A., and Serteyn, D. (2012). Physical fitness and mitochondrial respiratory capacity in horse skeletal muscle. *PLoS One* 7, e34890. doi:10.1371/journal.pone.0034890
- Wang, F., Wang, K., Xu, W., Zhao, S., Ye, D., Wang, Y., et al. (2017). SIRT5 desuccinylates and activates pyruvate kinase M2 to block macrophage IL-1 $\beta$  production and to prevent DSS-induced colitis in mice. *Cell Rep* 19, 2331–2344. doi:10.1016/j.celrep.2017.05.065
- Wong, M. L., Inserra, A., Lewis, M. D., Mastronardi, C. A., Leong, L., Choo, J., et al. (2016). Inflammation signaling affects anxiety- and depressive-like behavior and gut microbiome composition. *Mol. Psychiatry* 21, 797–805. doi:10.1038/mp.2016.46
- Wu, Z., Huang, X., Feng, Y., Handschin, C., Feng, Y., Gullicksen, P. S., et al. (2006). Transducer of regulated CREB-binding proteins (TORCs) induce PGC-1 $\alpha$  transcription and mitochondrial biogenesis in muscle cells. *Proc. Natl. Acad. Sci. USA* 103, 14379–14384. doi:10.1073/pnas.0606714103
- Yardeni, T., Tanes, C. E., Bittinger, K., Mattei, L. M., Schaefer, P. M., Singh, L. N., et al. (2019). Host mitochondria influence gut microbiome diversity: a role for ROS. *Sci. Signal.* 12, eaaw3159. doi:10.1126/scisignal.aaw3159
- Zhang, Y., Kumarasamy, S., Mell, B., Cheng, X., Morgan, E. E., Britton, S. L., et al. (2020). Vertical selection for nuclear and mitochondrial genomes shapes gut microbiota and modifies risks for complex diseases. *Physiol. Genomics* 52, 1–14. doi:10.1152/physiolgenomics.00089.2019
- Zheng, C., Zhang, S., Ragg, S., Rafferty, D., and Vitek, O. (2011). Identification and quantification of metabolites in  $^1\text{H}$  NMR spectra by Bayesian model selection. *Bioinformatics* 27, 1637–1644. doi:10.1093/bioinformatics/btr118

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

Copyright © 2021 Mach, Moroldo, Rau, Lecardonnel, Le Moyec, Robert and Barrey. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.