Electronic Supplementary Material (ESI) for Analyst.

Supplemental Information

Integration of metabolomics and proteomics to reveal metabolic characteristics of High Intensity Interval Training

Jingjing Zhao, Yang Wang, Dan Zhao, Lizhen Zhang, Peijie Chen* and Xin Xu*

Shanghai anti-doping laboratory, Shanghai University of Sport, Changhai Road 399, Shanghai, 200438, China.

E-mail: chenpeijie@sus.edu.cn; xxu2000@outlook.com; Fax: +86-21-65506702; Tel: +86-21-65506702

List of Supplemental Information

Experimental section containing subjects and baseline testing, experimental design, chemicals and reagents, establishment of the spectral library by DDA mode as well as statistical analysis.

Table S1. Descriptive characteristics of participants.

Fig. S1 Superimposed LC-MS total ion chromatography (TIC) of QC data obtained in negative ion mode.

Fig. S2 Extracted ion chromatogram (EIC) of creatinine of QC data in the negative ion mode. In this mode, the relative standard deviation (RSD) of retention time and the peak area were less than 5%.

Fig. S3 Heatmap visualization constructed based on the differential metabolic pathways compared post-exercise to pre-exercise.

Fig. S4 Heatmap visualization constructed based on the differential metabolic pathways compared the recovery to post-exercise.

Fig. S5 (a) PLS-DA plot of the proteome data of pre-exercise and post-exercise. (b) PLS-DA plot of the proteome data of post-exercise and the recovery.

Fig. S6 (a) Volcano plots of the proteome data compared post-exercise to pre-exercise. (b) Volcano plots of the proteome data compared the recovery to post-exercise.

Fig. S7 Heatmap visualization constructed based on the differential proteins compared post-exercise to pre-exercise.

Fig. S8 Heatmap visualization constructed based on the differential proteins compared the recovery to post-exercise.

Fig. S9 KEGG (kyoto encyclopedia of genes and genomes) pathway plot of the differential proteins compared post-exercise to pre-exercise (a) and compared the recovery to post-exercise (b).

Fig. S10 Protein-protein interaction (PPI) network analysis of the 102 significant proteins compared post-exercise to pre-exercise. The proteins were subjected to the STRING database. Every circular

node represents a protein and edge represents protein-protein associations with known, predicted or other sorts of interactions. Proteins contained in the blue cycle were associated to immunologic function.

Fig. S11 Protein-protein interaction (PPI) network analysis of the 134 significant proteins compared the recovery to post-exercise. The proteins were subjected to the STRING database. Every circular node represents a protein and edge represents protein-protein associations with known, predicted or other sorts of interactions. Proteins contained in the blue cycle were associated to immunologic function. Proteins in the red cycle participated in glycolysis were relevant to the energy metabolism.

The differential metabolites identified among pre-exercise, post-exercise and the recovery were listed in sheet 1 of Supplemental Dataset.

The differential proteins identified between pre-exercise and post-exercise were listed in sheet 2 of Supplemental Dataset.

The differential proteins identified between post-exercise and the recovery were listed in sheet 3 of Supplemental Dataset.

1. Experimental Section

1.1 Subjects and Baseline Testing

Twenty-three healthy players from Under-19 Youth Team of a professional soccer club volunteered to take part in this study. All participants had a professional training period of 5-10 years, and an intensive training experience of more than 3 years. Measurement of individualized VO_2 max was performed on a motorized treadmill (Cosmed, Rome, Italy). Heart rate were measured by a portable heart-rate monitor (POLAR Team Pro, Finland). Body weight and height were measured using an ultrasound height and weight measuring instrument (HYSTL200, China). Body mass index (BMI) was calculated as body mass divided by squared height. Descriptive characteristics of the subjects are summarized by means and standard deviations in Table S1. There were no significant di□erences in age, anthropometric parameters and physical fitness at baseline among the subjects.

1.2 Experimental Design

All human subjects provided informed consent before participating in this study and all performance data and biochemical data were anonymized. The study was approved by the ethics committee of Shanghai University of Sport, China (NO. 102772019RT050). The Yo-Yo intermittent recovery (YYIR) was used as HIIT model in the study. In YYIR test, the subjects underwent a repeated 2×20-m shuttle runs between the starting, turning, and finishing line at a progressively increased speed. Between each running bout, the subjects had a 10-s active rest period, consisting of 2×5 m of jogging. When the subjects had failed twice to reach the finishing line in time, the distance covered is recorded and end the exercise.

In order to research the influence of HIIT on athlete performance, the subjects underwent a two-day trial. Urine samples were collected respectively at pre-exercise in the first morning (labelled as Pre, first pass urine), 30 min post-exercise (labelled as Post) and 18 hours of recovery (labelled as Rec, first pass urine in the second morning). In the last, a total of 69 urine samples were collected, aliquoted and stored at -80 °C until analysis.

1.3 Chemicals and Reagents

HPLC grade acetonitrile (ACN), tris-(2-Carboxyethyl) phosphine, hydrochloride (TCEP) and the protease inhibitor cocktail were obtained from Thermo Fisher Scientific, UK. HPLC grade Methanol, ammonium formate, formic acid, sodium dodecyl sulfate (SDS), iodoacetamide (IAM), triethylammonium bicarbonate (TEAB), urea, acetone and ammonium hydroxide solution (28-30%) were purchased from Sigma-Aldrich, UK. Other chemicals, except as otherwise noted, were all analytical grade. Deionized water purified from Millipore (18 MΩ·cm) was used for all dilutions.

1.4 Establishment of the spectral library by DDA mode

The peptide mixture of thirty samples was re-dissovled in the buffer A (buffer A: 20 mM ammonium formate in water, pH 10.0, adjusted with ammonium hydroxide), and then fractionated by high pH separation using Ultimate 3000 system (ThermoFisher scientific, MA, USA) connected to a reverse phase XBridge C18 column (4.6 mm × 250 mm, 5 μm, Waters, MA, USA). High pH separation was performed using a linear gradient, starting from 5% B to 45% B in 40 min (B: 20mM ammonium formate in 80% ACN, pH 10.0, adjusted with ammonium hydroxide). The column was re-

equilibrated at the initial condition for 15 min. The column flow rate was maintained at 1 mL/min and the column temperature was maintained at 30℃. Six fractions were collected; each fraction was dried in a vacuum concentrator for the next step.

The peptides were re-dissolved in solvent A (A: 0.1% formic acid in water) and analyzed by on-line nanospray LC-MS/MS on an Orbitrap Fusion Lumos coupled to EASY-nLC 1200 system (Thermo Fisher Scientific, MA, USA). 2 μL peptide sample was loaded on analytical column (Acclaim PepMap C18, 75 μm × 25 cm) and separated with 60-min gradient, from 8% to 36% B (B: 0.1% formic acid in 80% ACN). The column flow rate was maintained at 600 nL/min. The electrospray voltage of 2 kV versus the inlet of the mass spectrometer was used.

The mass spectrometer was run under data dependent acquisition mode, and automatically switched between MS and MS/MS mode. The parameters was: (1) MS: scan range (m/z)=350–1800; resolution=60,000; AGC target=4e5; maximum injection time=50 ms; dynamic exclusion=40s; (2) HCD-MS/MS: resolution=15,000; AGC target=5e4; maximum injection time=30 ms; collision energy=30.

Raw Data of DDA were processed and analyzed by Spectronaut 13 (Biognosys AG, Switzerland) with default settings to generate an initial target list. Spectronaut was set up to search the database of Homo_sapiens_2017 (20414 entries) and BSA assuming trypsin as the digestion enzyme. Carbamidomethyl (C) was specified as the fixed modification. Oxidation (M) was specified as the variable modifications. Qvalue (FDR) cutoff on precursor and protein level was applied 1%.

1.5 Statistical analysis

T test was used to compare peak areas of metabolites. A two-way mixed analysis of variance (ANOVA) with repeated measures was used to test the main effect of pre-exercise, post-exercise and the recovery on the outcome variables. All experimental data are presented as means ± standard deviation (SD). A p value of < 0.05 was used as the criterion for statistical significance. Statistical significance was accepted at $*p < 0.05$, $**p < 0.01$, $***p < 0.001$, $****p < 0.0001$.

Multivariable statistical analysis of metabolome and proteome data was processed using Metaboanalyst 4.0 (McGill University, Quebec, Canada). Partial least squares regression-discriminant analysis (PLS-DA), a supervised multivariate regression technique, was performed to identify components that best differentiate between predefined classes of samples while dissecting orthogonal components which do not differentiate between these classes. In this study, PLS-DA models were established to visualize metabolic differences among pre-exercise (Pre), immediately post-exercise (Post) and the recovery after exercise (Rec).

Table S1. Descriptive characteristics of participants.

Total(n=23)	Mean ± SEM
Age (years)	17±2
Training age (years)	7.8 ± 2.7
Hight (cm)	175.34 ± 4.15
Weight (kg)	64.72 ± 7.13
BMI (kg/m2)	23.9 ± 0.6
Resting heart rate (bpm)	65.37 ± 8.45
Peak heart rate (bpm)	198.56 ± 5.69
Estimate peak VO_2 (mL/min/kg)	58.4 ± 5.8

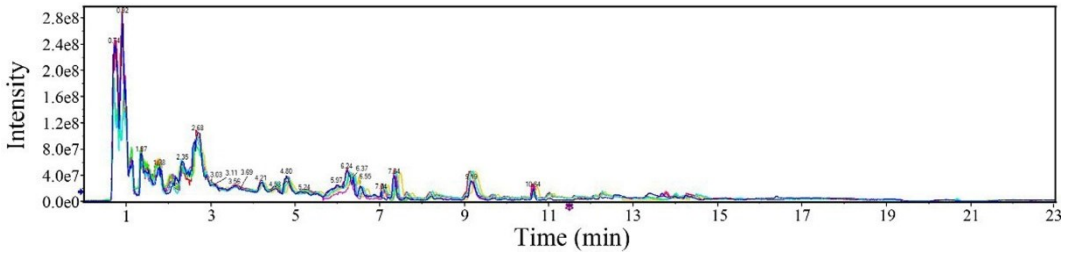

Fig. S1 Superimposed LC-MS total ion chromatography (TIC) of QC data obtained in negative ion mode.

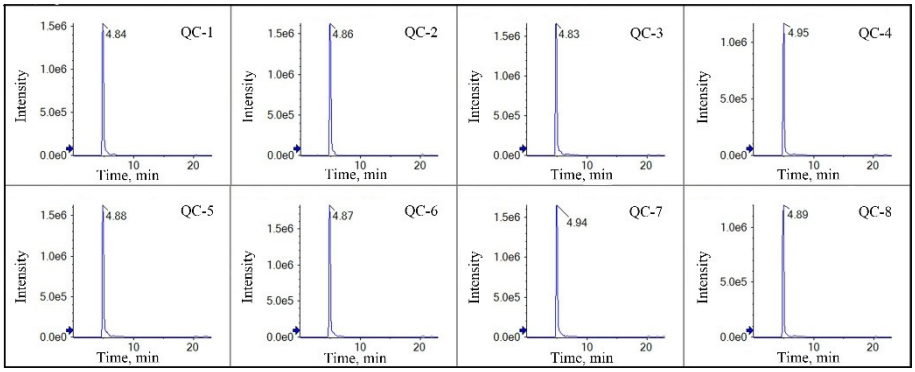

Fig. S2 Extracted ion chromatogram (EIC) of creatinine of QC data in the negative ion mode. In this mode, the relative standard deviation (RSD) of retention time and the peak area were less than 5%.

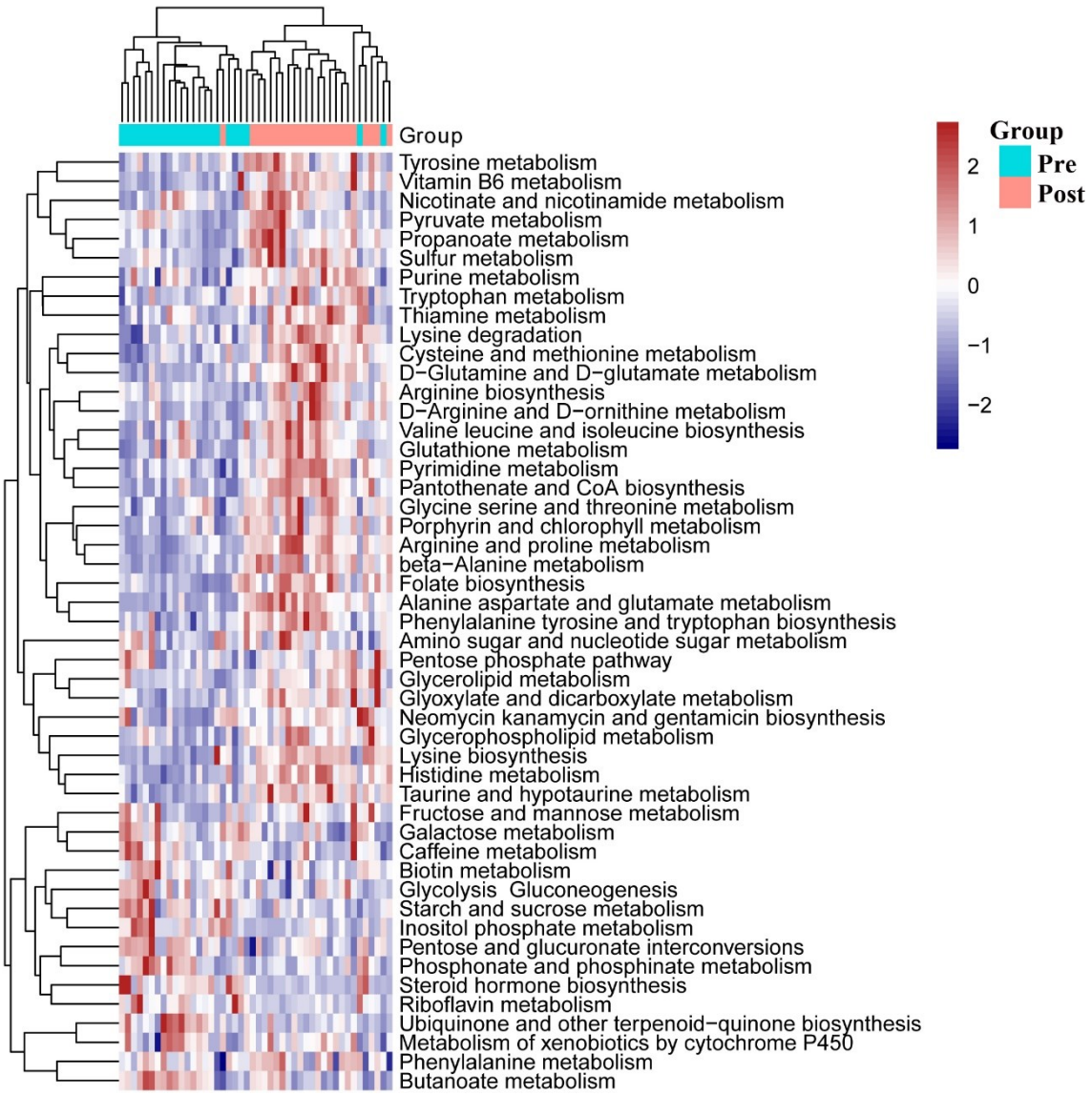

Fig. S3 Heatmap visualization constructed based on the differential metabolic pathways compared post-exercise to pre-exercise.

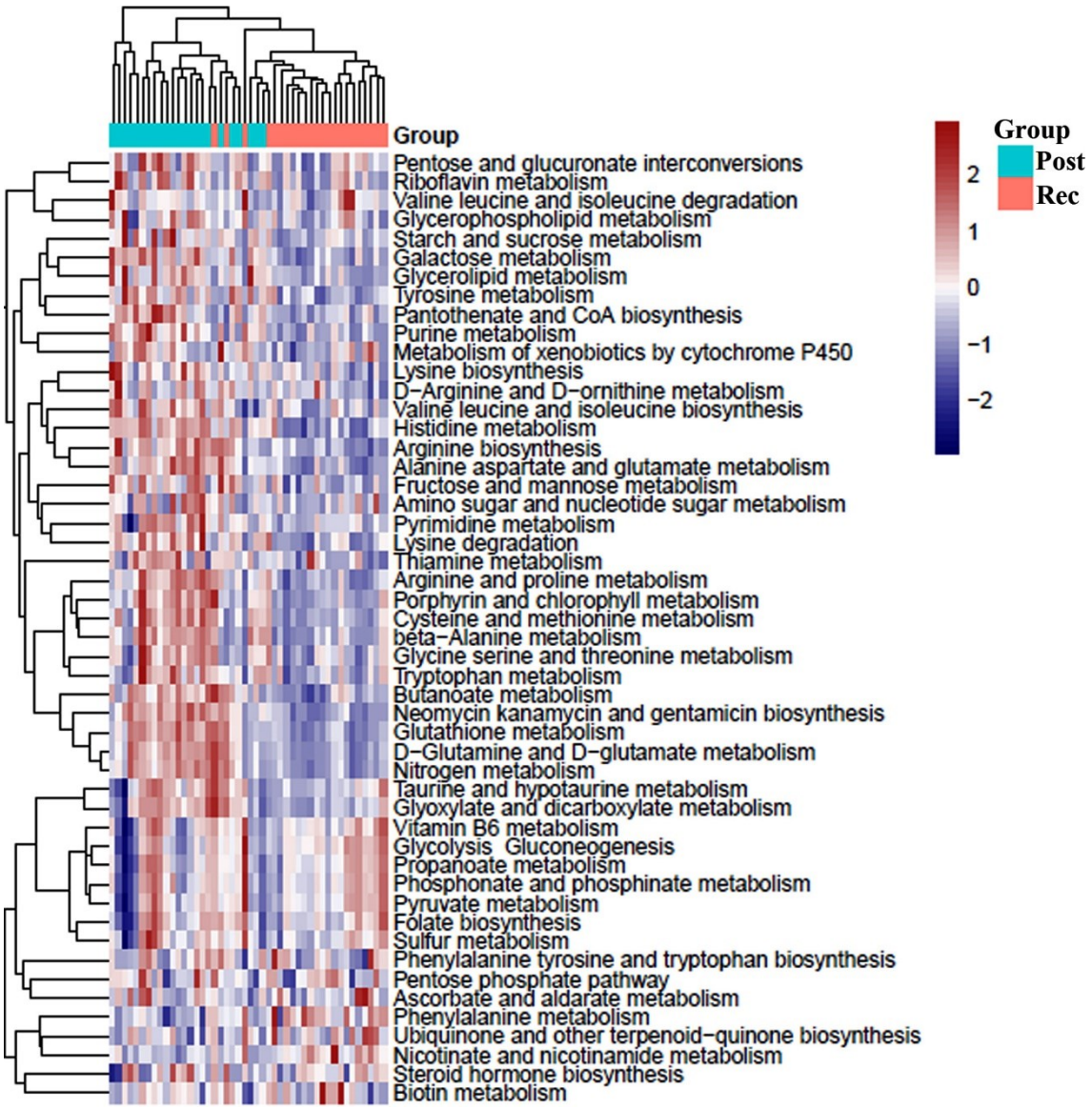

Fig. S4 Heatmap visualization constructed based on the differential metabolic pathways compared the recovery to post-exercise.

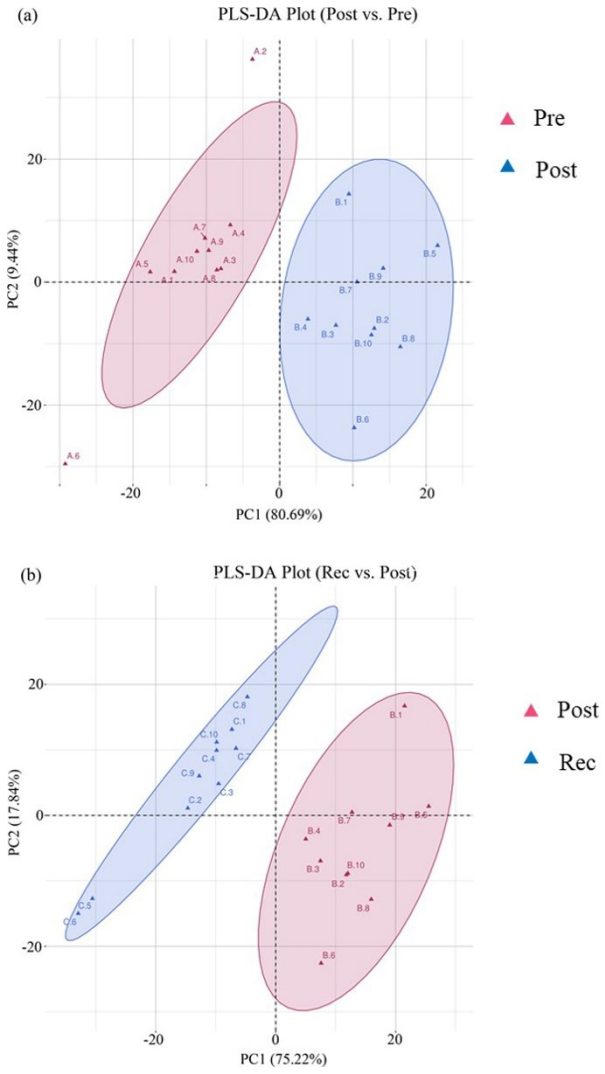

Fig. S5 (a) PLS-DA plot of the proteome data of pre-exercise and post-exercise. (b) PLS-DA plot of the proteome data of post-exercise and the recovery.

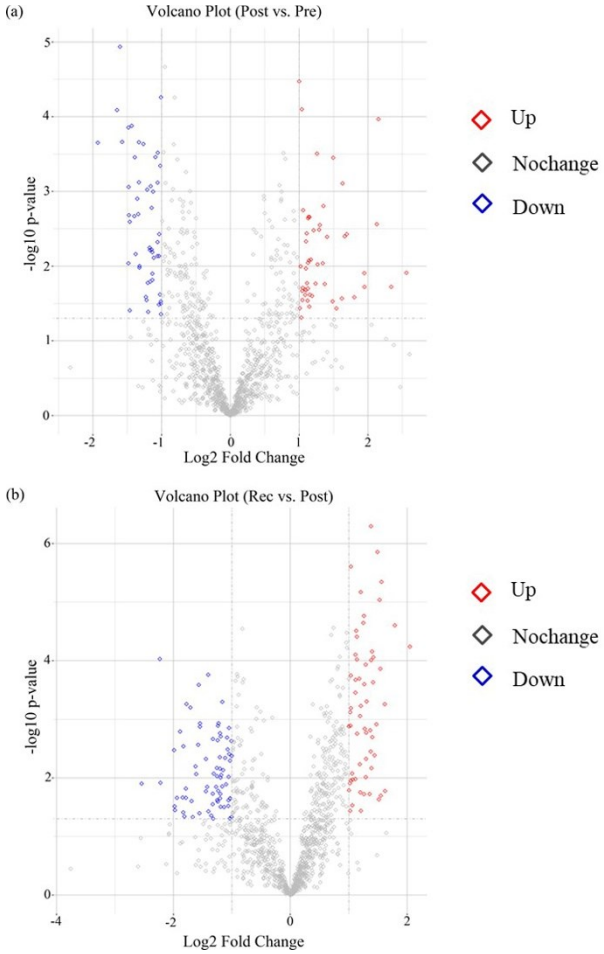

Fig. S6 (a) Volcano plots of the proteome data compared post-exercise to pre-exercise. (b) Volcano plots of the proteome data compared the recovery to post-exercise.

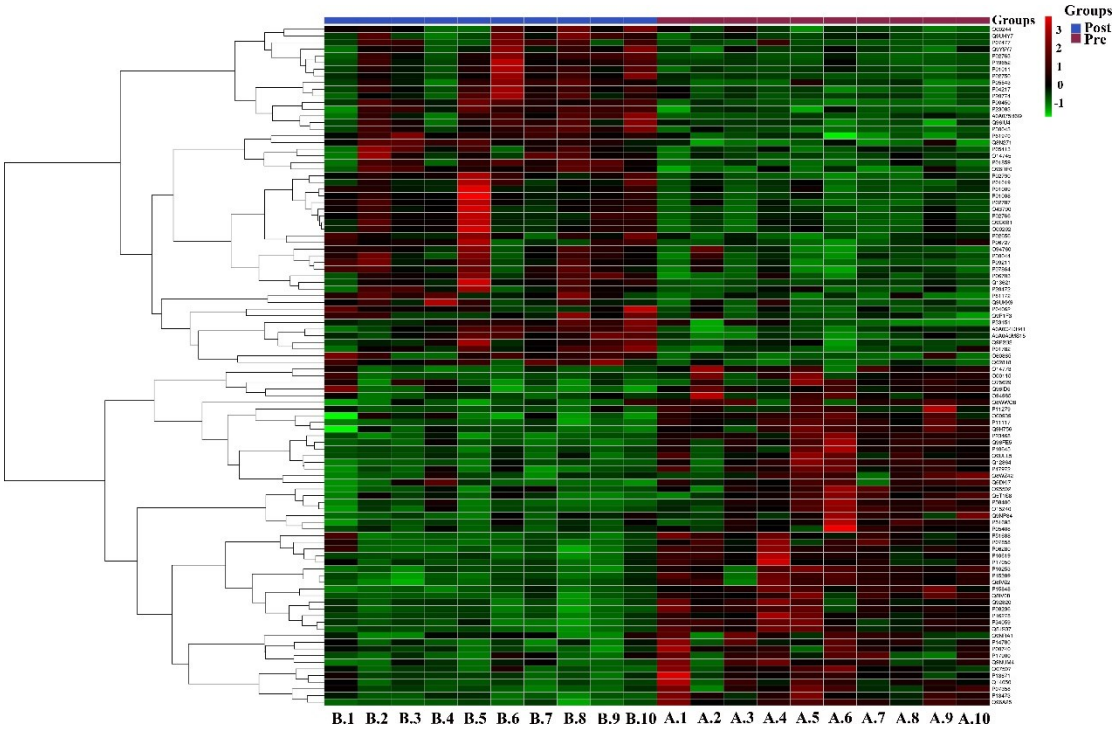

Fig. S7 Heatmap visualization constructed based on the differential proteins compared post-exercise to pre-exercise.

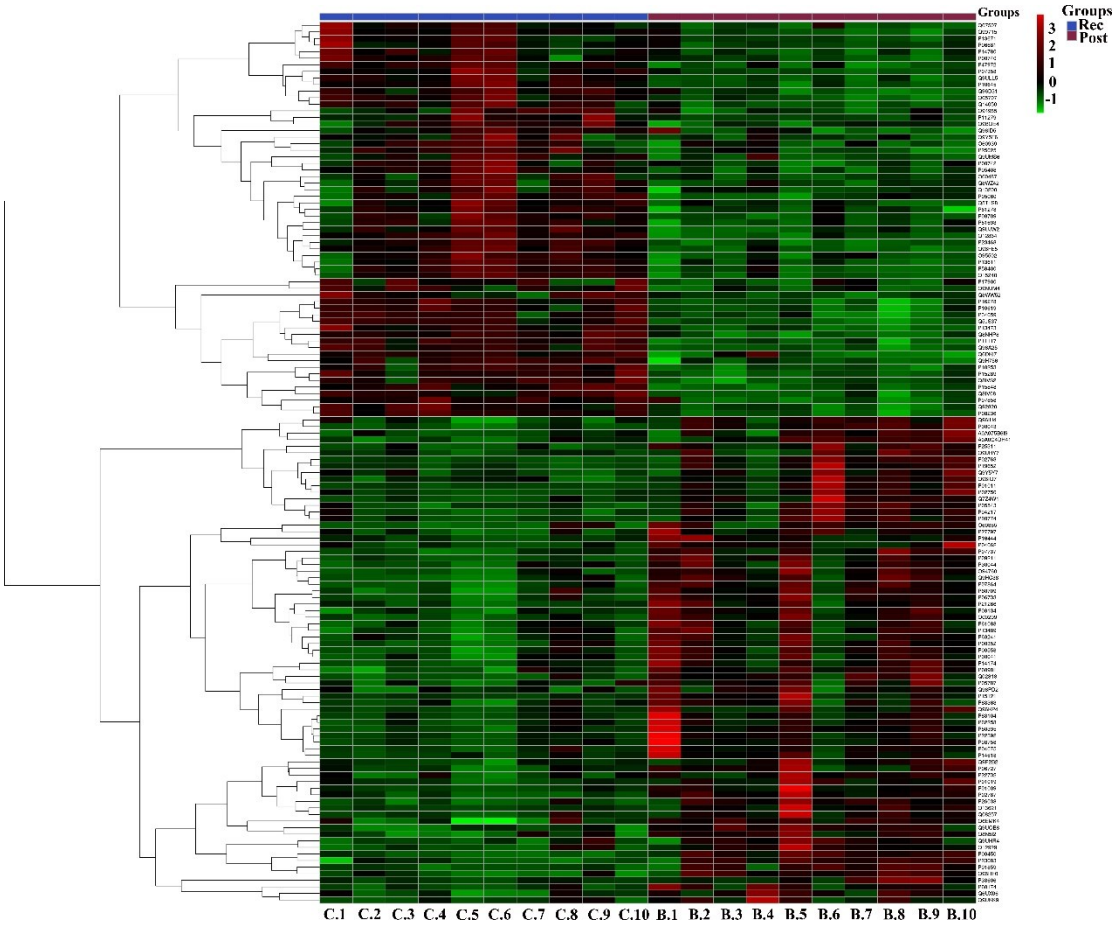

Fig. S8 Heatmap visualization constructed based on the differential proteins compared the recovery to post-exercise.

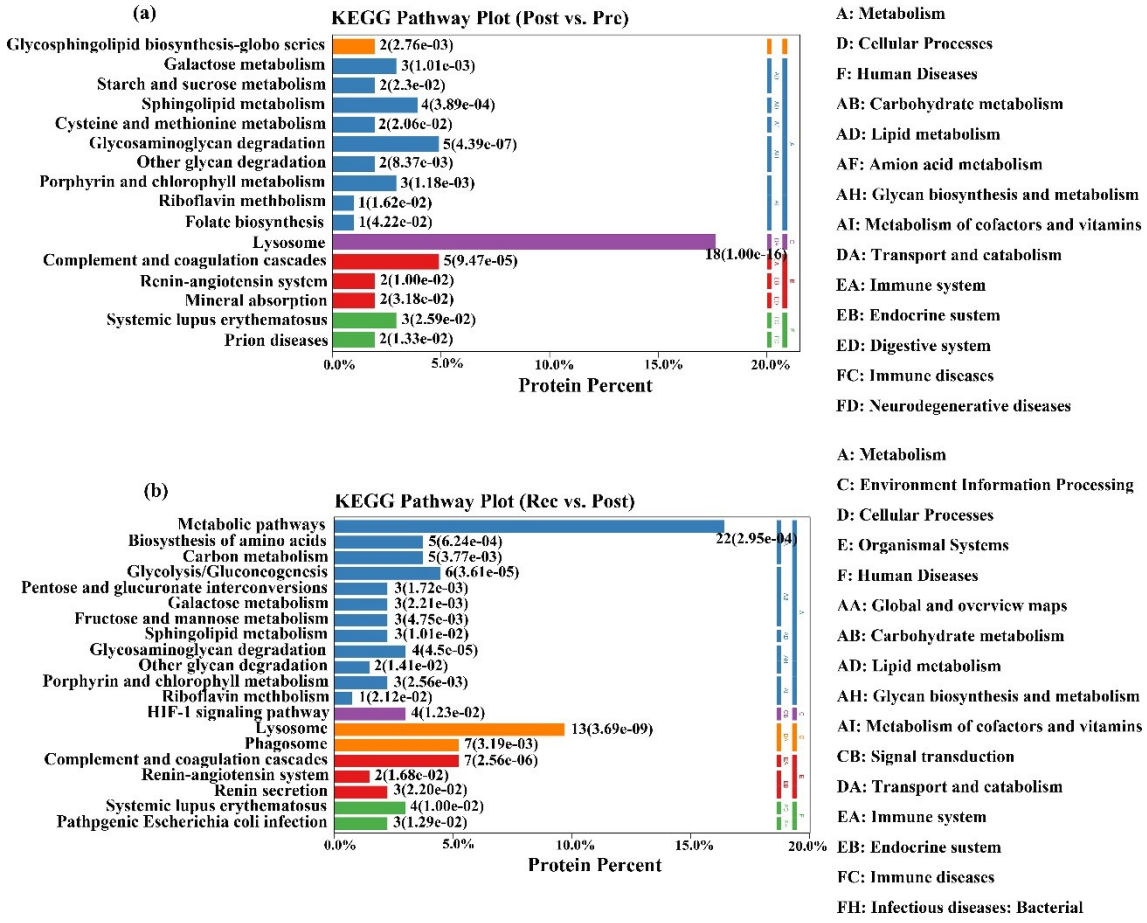

Fig. S9 KEGG (kyoto encyclopedia of genes and genomes) pathway plot of the differential proteins compared post-exercise to pre-exercise (a) and compared the recovery to post-exercise (b).

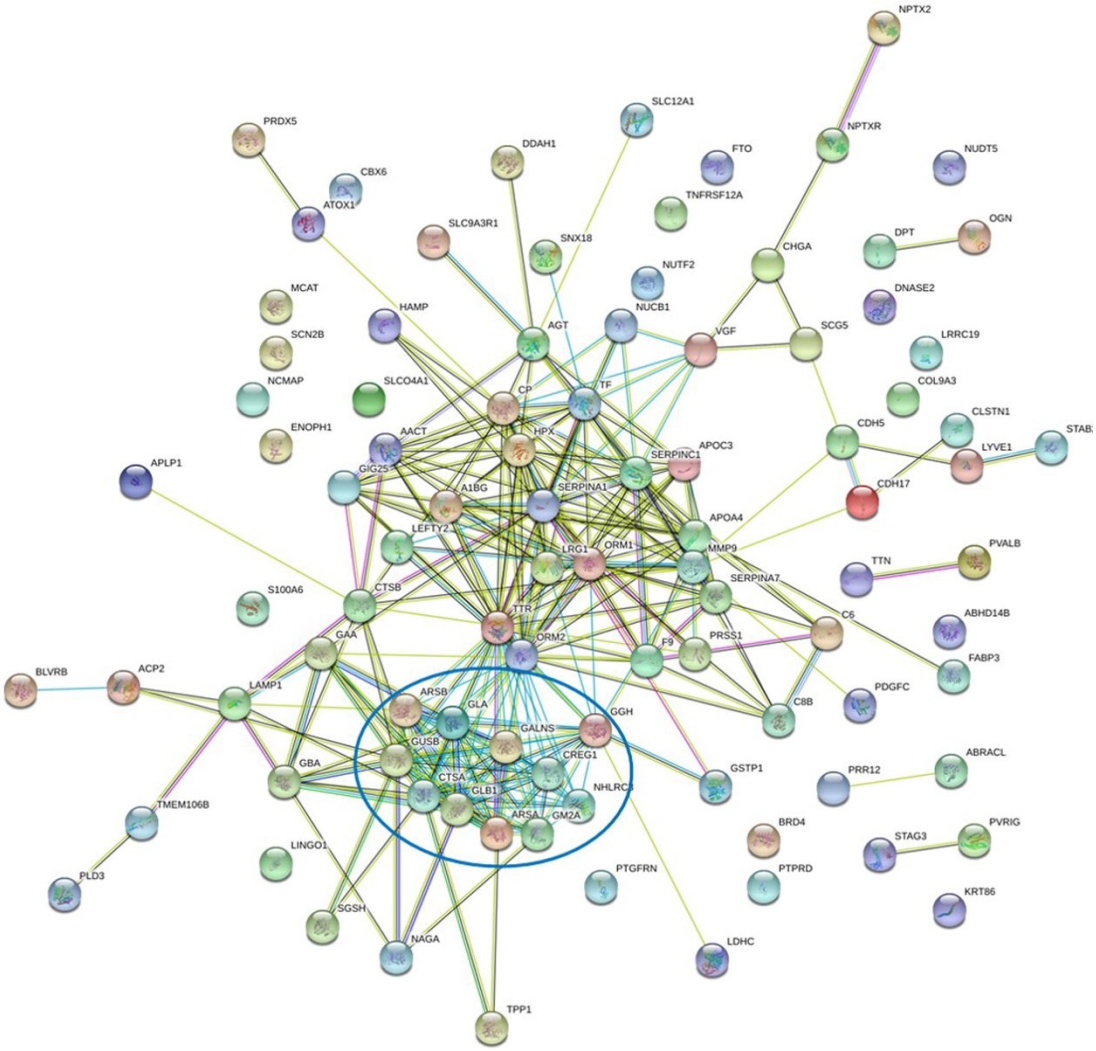

Fig. S10 Protein-protein interaction (PPI) network analysis of the 102 significant proteins compared post-exercise to pre-exercise. The proteins were subjected to the STRING database. Every circular node represents a protein and edge represents protein-protein associations with known, predicted or other sorts of interactions. Proteins contained in the blue cycle were associated to immunologic function.

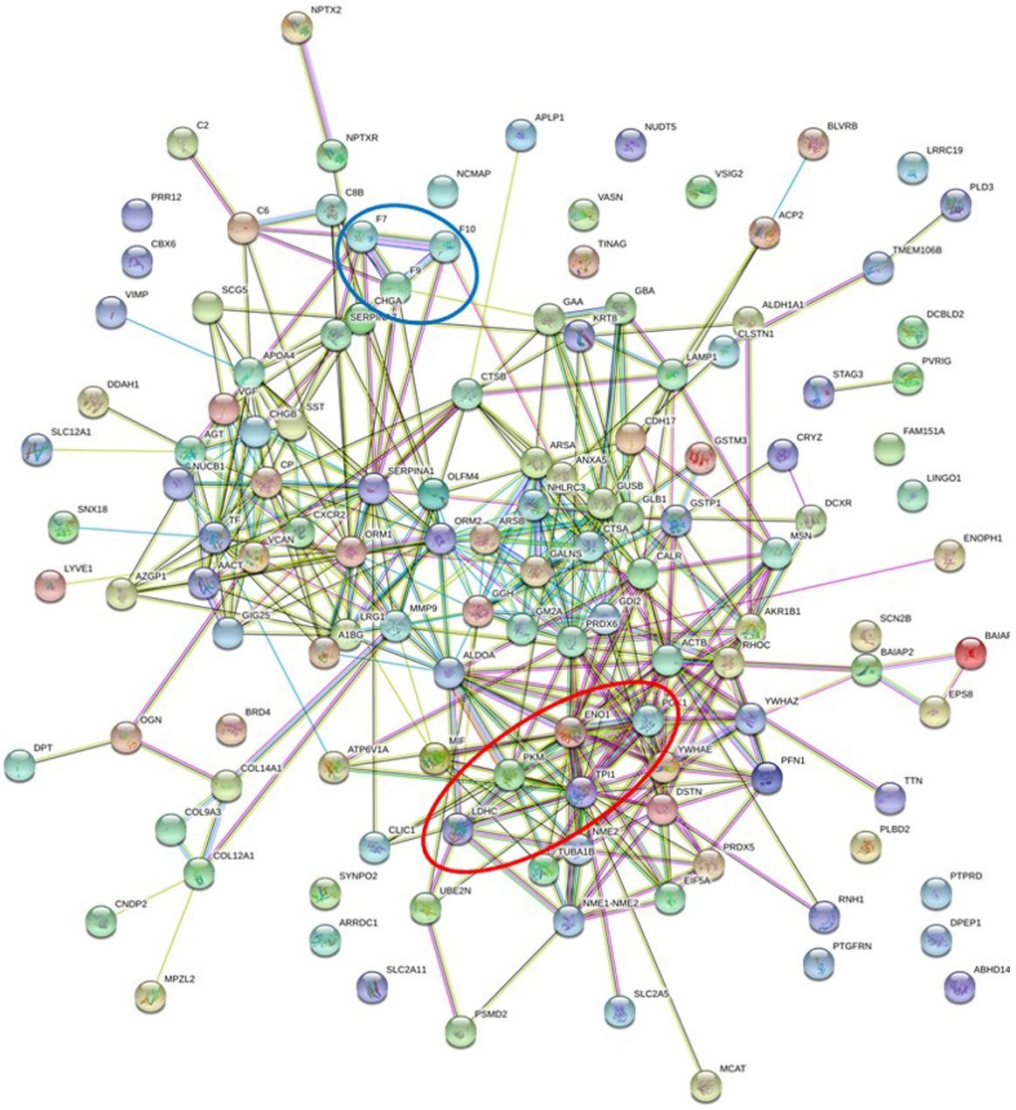

Fig. S11 Protein-protein interaction (PPI) network analysis of the 134 significant proteins compared the recovery to post-exercise. The proteins were subjected to the STRING database. Every circular node represents a protein and edge represents protein-protein associations with known, predicted or other sorts of interactions. Proteins contained in the blue cycle were associated to immunologic function. Proteins in the red cycle participated in glycolysis were relevant to the energy metabolism.