

Metabolic profiling of elite athletes with different cardiovascular demand

Fatima Al-Khelaifi^{1,2†}, Noha Yousri^{3†}, Francesco Donati⁴, Francesco Botrè⁴, Aishah Latiff¹, David Abraham², Aroon Hingorani², Costas Georgakopoulos¹, Karsten Suhre³, Mohamed A Elrayess^{1,2*}

¹ Anti Doping Laboratory Qatar, Sports City, Doha, Qatar.

² UCL-Medical School, Royal Free Campus, London NW3 2PF.

³ Department of Physiology and Biophysics, Weill Cornell Medical College in Qatar, Qatar-Foundation, P.O. Box 24144, Doha, Qatar.

⁴ Laboratorio Antidoping, Federazione Medico Sportiva Italiana, Largo Giulio Onesti 1, 00197 Rome, Italy.

† Both authors contributed equally to this work.

* Corresponding authors

Mohamed A Elrayess

ADLQ, Sports City, P.O Box 27775, Doha, Qatar

Office: (974) 4413 2845

Mobile: (974)33506764

Email: melrayess@adlqatar.qa

Abstract

Introduction: Intensive exercise of elite athletes can lead to physiological changes in the cardiovascular system in response to increased stroke volume and blood pressure, known collectively as cardiovascular demand (CD). This study aims to compare metabolic differences in elite athletes with high versus low-moderate CD and to identify the potential metabolic pathways underlying these differences.

Methods: Metabolic profiling of serum samples from 495 elite athletes from different sports disciplines (118 high CD and 377 low-moderate CD athletes) who participated in national or international sports events and tested negative for doping abuse at anti-doping laboratories was performed using non-targeted metabolomics-based mass spectroscopy combined with ultrahigh-performance liquid chromatography. Multivariate analysis was conducted using orthogonal partial least squares discriminant analysis. Differences in metabolic levels between high and low-moderate CD were assessed by univariate linear models. Gaussian graphical modelling (GGM) was constructed to identify metabolic networks based on their partial correlations.

Results: Metabolomics analysis revealed 112 novel metabolites that changed significantly with increased CD. Whereas diacylglycerols (DAGs) containing oleic acid were higher in low-moderate CD, DAGs containing arachidonic were enriched in high CD together with branched chain amino acids, plasminogens, phosphatidylcholines (PC) and phosphatidylethanolamines (PE), potentially marking increased risk of cardiovascular disease in the high CD group. Gamma glutamyl amino acids and glutathione metabolism were increased in low-moderate CD group, suggesting more efficient oxidative stress scavenging mechanisms than the high CD group. Additionally, various monohydroxy fatty acids and acyl-carnitines were too increased in low-moderate CD, reflecting increased fatty acid oxidation. GGM sub-networks identified 6 subnetworks that captured the major metabolic pathways perturbed in relation to differences in CD including fatty acids beta oxidation and oxidative stress.

Conclusion: This data provides evidence that athletes with different CD exhibit a distinct metabolic profile that reflects energy generation and oxidative stress and potentially places the high CD group at higher risk of cardiovascular disease. Replication studies are warranted to confirm these metabolic differences in independent data sets, aiming ultimately for identification of individual biomarkers for assessing health, performance and recovery of elite athletes with different CD.

Introduction

Athletes competing in national or international sport events are considered elite athletes [1]. Their intensive physical training is associated with electrical, structural and functional myocardial adaptations that help improve their sport's performance [2, 3]. The type and intensity of their respective sport disciplines determine the metabolic demands for their local and systemic blood flow, whereas the duration of training causes regulatory disturbances and constraints in their cardiovascular function over time [4]. In response to aerobic exercise, for example, the cardiovascular system could reach up to 80% of maximum cardiac output in order to meet the need to deliver oxygen to the exercising body organs while regulating the mean arterial pressure [5, 6].

Information related to peak static (maximal voluntary contraction, MVC) and dynamic (maximal oxygen uptake, MaxO₂) components achieved during competition for athletes from different sporting disciplines are used to classify them into various classes of cardiovascular demand [7]. The increasing dynamic component results in a greater cardiac output whereas the increasing static component results in an elevated blood pressure load. Accordingly, total cardiovascular demand (CD) that comprises both cardiac output and blood pressure is used to classify sports into five sport groups that range from low static/low dynamic (such as golf and rifle) to high static/high dynamic (such as boxing and cycling) [7]. Despite multiple physiological studies describing cardiac output and blood pressure of different groups of athletes [8], a more comprehensive monitoring of blood metabolic biomarkers may provide valuable insight into the physiological and pathophysiological alterations underlying athletes performance and health.

Metabolomics presents a comprehensive approach for detecting metabolic changes in response to dietary, lifestyle and environmental factors [9], including profiles associated with performance, fatigue and health issues [10, 11]. Previous non-targeted metabolomics studies revealed metabolic alterations in response to exercise [10, 12, 13], including changes in glucose, lipid, amino acid and energy metabolites [10, 13], marking adenosine triphosphate (ATP) synthesis and beta-oxidation of fatty acids [12] as well as elevation in plasma lactate [14, 15] and adenine breakdown products [16]. Components of tricarboxylic acid (TCA) cycle were also changed in response to aerobic energy production in skeletal muscle biopsies [17, 18]. We have recently shown in a pilot study of 191 elite athletes that high-power and high-endurance elite sports exhibit a distinct metabolic profile that reflects steroid biosynthesis, fatty acid metabolism, oxidative stress and energy-related metabolites [19].

Despite multiple studies focusing on the impact of exercise on athletes' metabolomics, profiling of metabolic changes in elite athletes from various sport disciplines characterized by different CD would provide deeper insight into their physiological state and cardiovascular adaptation in response to their respective sports [11]. Assessment of these changes would provide valuable measures of the current physical status of the athletes and their adaptation to training. This could help in directing their future training programs and preventing potential disorders associated with excessive exercise as well as improving their overall performance. This study aims to utilize a non-targeted metabolomics profiling to identify changes in metabolites with increased CD in athletes belonging to different sport disciplines.

Methods

Study design

Four hundred ninety six elite athletes (87% males, 13% females) from different sports disciplines, who participated in national or international sports events and tested negative for doping substances at anti-doping laboratories in Qatar and Italy, were included in this study in accordance with the World Medical Association Declaration of Helsinki. All protocols were approved by the Institutional Research Board of anti-doping lab Qatar (F2014000009). There was no evidence of population stratification in sport groups based on athletes' ethnicities (85% Europeans, 10% Americans and 5% Africans). Spare serum samples, collected for anti-doping human growth hormone tests, were used for metabolomics studies as described previously [19]. Accordingly, only information related to sport type, ethnic group and gender were available for researchers. Total CD was used to classify sports into two groups (low-moderate CD (n=377), shown in white in Table 1, versus high CD (n=118), shown in grey in Table 1.

Table 1. Classification of study participants.

		Endurance	
		Moderate (40-70% VO ₂ max)	High (>70% VO ₂ max) (n=63)
Power	High (>50% MVC)	Wrestling (3M), Judo (3M)	Boxing (1M/16F), Heptathlon (1M), Rowing (6M/7F), Cycling (31M/4F)
	Moderate (20-50% MVC)	Athletics (15M/22F), Rugby (16M), Triple Jump (1M)	Athletics 200-800m (4M), Hockey (1F), Skiing Cross Country (1M), Basketball (3M), Swimming (22M/16F)
	Low (<20% MVC)	Baseball (2M), Volleyball (1M)	Tennis (1M/1F), Soccer (315M), Athletics 1500-3000m (3M)

Distribution of elite athletes in various categories based on sport type-associated peak dynamic (maximal oxygen uptake percentage; VO₂) and peak static (maximal voluntary muscle contraction percentage; MVC) components achieved during competition as described previously [7]. The number and gender (M for males and F for females) of participants in each group are also indicated. Sports types were further categorized into two classes based on CD (high in grey and low in white) that were used in the statistical analysis.

Metabolomics

Metabolomics profiling was performed using validated protocols at Metabolon, Durham, NC, USA, which utilized Waters ACQUITY ultra-performance liquid chromatography (UPLC) and a Thermo Scientific Q-Exactive high resolution/accurate mass spectrometer interfaced with a heated electrospray ionization (HESI-II) source and Orbitrap mass analyzer operated at 35,000 mass resolution.

Sample preparation

Upon delivery, sample inventories were confirmed and samples were stored at -80°C until time of processing. Sample preparation was performed using automated MicroLab STAR® system (Hamilton Company). Recovery standards were added prior to the first step in the extraction process for QC purposes. Proteins were precipitated in methanol under vigorous shaking for 2 min (Glen Mills GenoGrinder 2000), followed by centrifugation to remove proteins, dissociate small molecules that are either bound to proteins or trapped in the precipitated protein matrix and to recover chemically diverse metabolites. The resulting extract was divided into five fractions: two for analysis by two separate reverse phase (RP)/UPLC-MS/MS methods with positive ion mode electrospray ionization (ESI) (for detection of positive early and positive late fractions), one for analysis by RP/UPLC-MS/MS with negative ion mode ESI, one for analysis by HILIC/UPLC-MS/MS with negative ion mode ESI (for the detection of the polar compounds), and one sample was reserved for backup. Samples were placed briefly on a TurboVap® (Zymark) to remove the organic solvent. The sample extracts were stored overnight under nitrogen before preparation for analysis.

Quality control (QC)

Several controls were analyzed together with the tested samples. These included a pooled matrix generated by taking a small volume of each experimental sample served as a technical replicate throughout the data set. Extracted water samples were used as blanks. Additionally, a cocktail of QC standards that were carefully chosen not to interfere with the measurement of endogenous compounds were spiked into every analyzed sample, allowing instrument performance monitoring and aiding chromatographic alignment. Instrument variability was determined by calculating the median relative standard deviation (RSD) for the standards that were added to each sample prior to injection into the mass spectrometers. Overall process variability was determined by calculating the median RSD for all endogenous metabolites (i.e., non-instrument standards) present in 100% of the pooled matrix samples. Experimental samples were randomized across the platform and run with QC samples that were spaced evenly among the injections.

Ultrahigh Performance Liquid Chromatography-Tandem Mass Spectroscopy (UPLC-MS/MS)

The dried sample extract was reconstituted in solvents specific to each of the four methods. Each reconstitution solvent contained a series of standards at fixed concentrations to ensure injection and chromatographic consistency. One aliquot was analyzed using acidic positive ion conditions, chromatographically optimized for more hydrophilic compounds. In this method, the extract was

gradient eluted from a C18 column (Waters UPLC BEH C18-2.1x100 mm, 1.7 μ m) using water and methanol, containing 0.05% perfluoropentanoic acid (PFPA) and 0.1% formic acid (FA). Another aliquot was also analyzed using acidic positive ion conditions, however it was chromatographically optimized for more hydrophobic compounds. In this method, the extract was gradient eluted from the same mentioned C18 column using methanol, acetonitrile, water, 0.05% PFPA and 0.01% FA and was operated at an overall higher organic content. Another aliquot was analyzed using basic negative ion optimized conditions using a separate dedicated C18 column. The basic extracts were gradient eluted from the column using methanol and water, however with 6.5mM Ammonium Bicarbonate at pH 8. The fourth aliquot was analyzed via negative ionization following elution from a HILIC column (Waters UPLC BEH Amide 2.1x150 mm, 1.7 μ m) using a gradient consisting of water and acetonitrile with 10mM Ammonium Formate, pH 10.8. The MS analysis alternated between MS and data-dependent MS/MS scans using dynamic exclusion. The scan range varied slightly between methods but covered 70-1000 m/z . Raw data files are archived and extracted as described below.

Data extraction and compound identification

Raw data was extracted, peak-identified and QC processed using Metabolon's hardware and software. These systems are built on a web-service platform utilizing Microsoft's NET technologies, which run on high-performance application servers and fiber-channel storage arrays in clusters to provide active failover and load-balancing. Compounds were identified by comparison to library entries of purified standards or recurrent unknown entities. Metabolon maintains a library based on authenticated standards that contains the retention time/index (RI), mass to charge ratio (m/z), and chromatographic data (including MS/MS spectral data) on all molecules present in the library. Furthermore, biochemical identifications are based on three criteria: the correct retention time/index to the authentic standard, the correct m/z within 10ppm of the authentic standard and the correct fragmentation spectrum (MS/MS) to the standard. The MS/MS scores are based on a comparison of the ions present in the experimental spectrum to the ions present in the library spectrum. While there may be similarities between these molecules based on one of these factors, the use of all three data points can be utilized to distinguish and differentiate biochemicals. More than 3300 commercially available purified standard compounds have been acquired and registered into Laboratory Information Management System (LIMS) for analysis on all platforms for determination of their analytical characteristics. Metabolon data analysts use proprietary visualization and interpretation software to confirm the consistency of peak identification among the various samples. Library matches for each compound were checked for each sample and corrected if necessary. Metabolon classified known metabolites by "super-pathway", representing chemical classes, and "sub-pathway", corresponding to the specific role of a compound in metabolism, on the basis of the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways [20].

Statistical analysis of metabolomics data

Metabolomics data were log-transformed to ensure normality. Batch correction was performed by Metabolon by rescaling each metabolite's median to 1. Principle component analysis (PCA) was performed to obtain a global view of the data using a multivariate analysis. Orthogonal partial least square discriminant analysis (OPLS-DA) was used to detect components that best distinguish

between predefined groups of samples whilst separating the orthogonal components which do not differentiate between these groups. In this study, OPLS-DA was used to compare elite athletes with low-moderate versus high CD. Both PCA and OPLS-DA were run using SIMCA 14 with a threshold percentage of missing metabolite values of less than 50% (A default metabolite-wise metabolite missingness threshold). There was no evidence of population stratifications in sport groups. Linear models for association analysis were run using the R statistical package (version 2.14, www.r-project.org/). A model incorporating CD as a categorical variable with two levels (low-moderate & high) was used after correcting for covariates including gender, hemolysis levels (determined visually by Metabolon) and ethnic groups. A stringent Bonferroni level of significance of $p < 0.05/751 = 6.6 \times 10^{-5}$ was used to infer association. Gaussian graphical modelling (GGM) was used to identify correlated metabolites, leading potentially to unbiased reconstruction of metabolic reactions as previously reported [21].

Results

Non-targeted metabolomics was applied to compare the metabolic signature of elite athletes with low-moderate versus high CD. Analysis detected 751 known metabolites (Supplemental Table 1), among which 112 were found significantly different between the two groups ($p < 6.6 \times 10^{-5}$), including 57 lipids, 25 amino acids, 12 peptides, 8 xenobiotics, 5 nucleotides, 3 carbohydrates and 2 cofactors and vitamins (Table 2). Considering hits sub-pathways, predominant changes between the two CD groups included diacylglycerols, fatty acid metabolism (acyl Carnitine), gamma-glutamyl amino acid, lysopho and phospholipids among others.

Table 2. Metabolites differentiating between athletes belonging to low-moderate versus high cardiovascular demand groups (Bonferroni significance).

Metabolite	SUB_PATHWAY	Change	p-value	Bonferroni p value
Aspartate	Alanine and Aspartate Metabolism	-0.8	1.14E-11	8.58E-09
N-acetylglucosamine/n-acetylgalactosamine	Aminosugar Metabolism	-0.8	3.40E-12	2.55E-09
N-acetylneuraminate	Aminosugar Metabolism	-0.7	9.82E-09	7.38E-06
O-methylcatechol Sulfate	Benzoate Metabolism	-0.5	5.09E-05	3.82E-02
Carnitine	Carnitine Metabolism	0.5	2.68E-06	2.01E-03
Lactosyl-n-palmitoyl-sphingosine (D18:1/16:0)	Ceramides	-0.6	8.50E-07	6.39E-04
Glycosyl-n-(2-hydroxynervonoyl)-sphingosine	Ceramides	-0.5	9.16E-06	6.88E-03
4-hydroxychlorothalonil	Chemical	0.6	2.51E-07	1.89E-04
Cortisol	Corticosteroids	0.6	5.99E-08	4.50E-05
Creatine	Creatine Metabolism	0.5	3.07E-06	2.30E-03
Palmitoyl-oleoyl-glycerol (16:0/18:1) [2]	Diacylglycerol	-0.6	2.62E-06	1.97E-03
Linoleoyl-arachidonoyl-glycerol (18:2/20:4) [1]	Diacylglycerol	0.5	3.73E-06	2.80E-03
Oleoyl-oleoyl-glycerol (18:1/18:1) [1]	Diacylglycerol	-0.5	9.38E-06	7.05E-03
Palmitoyl-arachidonoyl-glycerol (16:0/20:4) [2]	Diacylglycerol	0.6	1.17E-05	8.78E-03
Oleoyl-linoleoyl-glycerol (18:1/18:2) [2]	Diacylglycerol	-0.5	1.38E-05	1.04E-02
Oleoyl-oleoyl-glycerol (18:1/18:1) [2]	Diacylglycerol	-0.5	1.72E-05	1.29E-02
Palmitoleoyl-arachidonoyl-glycerol (16:1/20:4) [2]	Diacylglycerol	0.6	3.82E-05	2.87E-02
Oleoyl-linoleoyl-glycerol (18:1/18:2) [1]	Diacylglycerol	-0.5	5.66E-05	4.25E-02
N-acetylcarnosine	Dipeptide Derivative	-0.4	3.52E-05	2.64E-02
Leukotriene B4	Eicosanoid	-0.6	9.75E-07	7.32E-04
5-hete	Eicosanoid	-0.6	4.21E-06	3.16E-03
Leukotriene B5	Eicosanoid	-0.5	5.71E-06	4.29E-03
12-hete	Eicosanoid	-0.5	6.10E-05	4.58E-02
N-oleoyltaurine	Endocannabinoid	-0.5	1.03E-05	7.72E-03

Oleoylcarnitine (C18:1)	Fatty Acid Metabolism(Acyl Carnitine)	-0.7	1.86E-09	1.39E-06
Ximenoylcarnitine (C26:1)	Fatty Acid Metabolism(Acyl Carnitine)	-0.6	1.21E-08	9.09E-06
Myristoleoylcarnitine (C14:1)	Fatty Acid Metabolism(Acyl Carnitine)	-0.6	1.35E-08	1.01E-05
Palmitoleoylcarnitine (C16:1)	Fatty Acid Metabolism(Acyl Carnitine)	-0.6	6.50E-08	4.88E-05
Eicosenoylcarnitine (C20:1)	Fatty Acid Metabolism(Acyl Carnitine)	-0.6	3.23E-07	2.43E-04
Laurylcarnitine (C12)	Fatty Acid Metabolism(Acyl Carnitine)	-0.5	7.16E-06	5.38E-03
Linoleoylcarnitine (C18:2)	Fatty Acid Metabolism(Acyl Carnitine)	-0.5	9.83E-06	7.38E-03
Decanoylcarnitine (C10)	Fatty Acid Metabolism(Acyl Carnitine)	-0.5	1.67E-05	1.25E-02
Adipoylcarnitine (C6-dc)	Fatty Acid Metabolism(Acyl Carnitine)	-0.5	2.25E-05	1.69E-02
14-hdohe/17-hdohe	Fatty Acid, Monohydroxy	-0.8	4.35E-13	3.27E-10
13-hode + 9-hode	Fatty Acid, Monohydroxy	-0.7	1.89E-10	1.42E-07
4-hdohe	Fatty Acid, Monohydroxy	-0.7	2.84E-08	2.14E-05
Carotene Diol (1)	Food Component/Plant	0.7	4.51E-10	3.38E-07
Carotene Diol (2)	Food Component/Plant	0.7	1.34E-09	1.01E-06
Ergothioneine	Food Component/Plant	0.6	1.96E-06	1.47E-03
Eugenol Sulfate	Food Component/Plant	-0.5	1.70E-05	1.28E-02
Gamma-glutamylglutamate	Gamma-glutamyl Amino Acid	-0.8	4.90E-13	3.68E-10
Gamma-glutamylphenylalanine	Gamma-glutamyl Amino Acid	-0.8	2.70E-11	2.03E-08
Gamma-glutamylvaline	Gamma-glutamyl Amino Acid	-0.6	4.41E-08	3.31E-05
Gamma-glutamyltyrosine	Gamma-glutamyl Amino Acid	-0.6	4.75E-08	3.57E-05
Gamma-glutamyltryptophan	Gamma-glutamyl Amino Acid	-0.6	3.95E-07	2.97E-04
Gamma-glutamylisoleucine	Gamma-glutamyl Amino Acid	-0.6	4.47E-07	3.36E-04
Gamma-glutamylleucine	Gamma-glutamyl Amino Acid	-0.5	2.50E-06	1.88E-03
Gamma-glutamylthreonine	Gamma-glutamyl Amino Acid	-0.5	4.05E-06	3.04E-03
Gamma-glutamyl-epsilon-lysine	Gamma-glutamyl Amino Acid	-0.5	9.10E-06	6.84E-03
Gamma-glutamylhistidine	Gamma-glutamyl Amino Acid	-0.5	3.55E-05	2.66E-02
Gamma-glutamylglutamine	Gamma-glutamyl Amino Acid	0.5	3.65E-05	2.74E-02
Glutamate	Glutamate Metabolism	-0.8	1.47E-13	1.11E-10
Glutamine	Glutamate Metabolism	0.6	1.80E-06	1.35E-03
Beta-citrylglutamate	Glutamate Metabolism	-0.5	3.09E-05	2.32E-02
5-oxoproline	Glutathione Metabolism	-0.9	2.19E-16	1.65E-13
Serine	Glycine, Serine and Threonine Metabolism	-0.6	1.94E-08	1.45E-05
N-acetylthreonine	Glycine, Serine and Threonine Metabolism	-0.6	1.24E-06	9.32E-04
N-acetyls erine	Glycine, Serine and Threonine Metabolism	-0.5	5.70E-06	4.28E-03
Glycerate	Glycolysis, Gluconeogenesis, and Pyruvate Metabolism	-0.6	1.78E-08	1.33E-05
Imidazole Lactate	Histidine Metabolism	0.8	6.22E-13	4.67E-10
1-methylimidazoleacetate	Histidine Metabolism	-0.5	9.57E-06	7.19E-03
3-methyl-2-oxobutyrate	Leucine, Isoleucine and Valine Metabolism	0.9	8.42E-16	6.33E-13
4-methyl-2-oxopentanoate	Leucine, Isoleucine and Valine Metabolism	0.9	1.27E-15	9.57E-13
3-methyl-2-oxovalerate	Leucine, Isoleucine and Valine Metabolism	0.8	6.96E-12	5.23E-09
Isovalerate	Leucine, Isoleucine and Valine Metabolism	-0.6	2.63E-07	1.97E-04
N6,n6,n6-trimethyllysine	Lysine Metabolism	-0.6	1.25E-06	9.39E-04
1-linolenoyl-gpc (18:3)	Lysophospholipid	0.7	6.38E-09	4.79E-06
1-palmitoyl-gpa (16:0)	Lysophospholipid	-0.6	1.05E-07	7.87E-05
1-arachidonoyl-gpa (20:4)	Lysophospholipid	-0.6	1.98E-07	1.48E-04
1-oleoyl-gpa (18:1)	Lysophospholipid	-0.6	2.37E-07	1.78E-04
1-linoleoyl-gpi (18:2)	Lysophospholipid	0.5	1.87E-05	1.40E-02
1-(1-enyl-palmitoyl)-gpc (P-16:0)	Lysoplasmalogen	-0.6	3.16E-08	2.37E-05
1-(1-enyl-oleoyl)-gpe (P-18:1)	Lysoplasmalogen	-0.5	1.50E-05	1.12E-02
Caproate (6:0)	Medium Chain Fatty Acid	-0.5	1.39E-05	1.04E-02
Methionine Sulfoxide	Methionine, Cysteine, SAM and Taurine Metabolism	-0.8	1.60E-13	1.20E-10
N-acetylmethionine Sulfoxide	Methionine, Cysteine, SAM and Taurine Metabolism	-0.8	2.73E-13	2.05E-10
N-formylmethionine	Methionine, Cysteine, SAM and Taurine Metabolism	0.6	5.90E-07	4.43E-04
Cysteine Sulfinic Acid	Methionine, Cysteine, SAM and Taurine Metabolism	-0.5	9.83E-06	7.39E-03

1-myristoyl-2-arachidonoyl-gpc (14:0/20:4)	Phosphatidylcholine (PC)	0.9	2.27E-15	1.70E-12
1-linoleoyl-2-arachidonoyl-gpc (18:2/20:4n6)	Phosphatidylcholine (PC)	0.8	1.88E-12	1.41E-09
1-myristoyl-2-palmitoyl-gpc (14:0/16:0)	Phosphatidylcholine (PC)	0.7	4.62E-10	3.47E-07
1-palmitoyl-2-palmitoleoyl-gpc (16:0/16:1)	Phosphatidylcholine (PC)	0.7	1.11E-09	8.35E-07
1-palmitoleoyl-2-linolenoyl-gpc (16:1/18:3)	Phosphatidylcholine (PC)	0.8	3.27E-07	2.46E-04
1-palmitoyl-2-arachidonoyl-gpc (16:0/20:4n6)	Phosphatidylcholine (PC)	0.6	9.99E-07	7.50E-04
1,2-dilinoleoyl-gpc (18:2/18:2)	Phosphatidylcholine (PC)	0.5	3.40E-05	2.55E-02
1-palmitoyl-2-arachidonoyl-gpc (16:0/20:4)	Phosphatidylethanolamine (PE)	0.6	9.38E-08	7.04E-05
1-stearoyl-2-arachidonoyl-gpc (18:0/20:4)	Phosphatidylethanolamine (PE)	0.5	9.06E-06	6.80E-03
1-palmitoyl-2-linoleoyl-gpc (16:0/18:2)	Phosphatidylethanolamine (PE)	0.5	1.07E-05	8.02E-03
1-palmitoyl-2-oleoyl-gpc (16:0/18:1)	Phosphatidylethanolamine (PE)	0.5	1.60E-05	1.20E-02
1-palmitoyl-2-docosahexaenoyl-gpc (16:0/22:6)	Phosphatidylethanolamine (PE)	0.5	3.77E-05	2.83E-02
Choline	Phospholipid Metabolism	-0.7	1.77E-09	1.33E-06
1-(1-enyl-palmitoyl)-2-oleoyl-gpc (P-16:0/18:1)	Plasmalogen	0.6	8.17E-08	6.14E-05
1-(1-enyl-stearoyl)-2-linoleoyl-gpc (P-18:0/18:2)	Plasmalogen	0.5	5.27E-06	3.96E-03
1-(1-enyl-palmitoyl)-2-arachidonoyl-gpc (P-16:0/20:4)	Plasmalogen	0.5	6.36E-05	4.78E-02
5alpha-pregnan-3beta,20beta-diol Monosulfate (1)	Progestin Steroids	-0.5	5.60E-05	4.20E-02
Allantoin	Purine Metabolism, (Hypo)Xanthine/Inosine	-0.5	3.05E-06	2.29E-03
N1-methyladenosine	Purine Metabolism, Adenine containing	0.6	9.81E-07	7.36E-04
Adenine	Purine Metabolism, Adenine containing	0.5	4.49E-06	3.37E-03
Orotate	Pyrimidine Metabolism, Orotate containing	-0.5	4.99E-06	3.75E-03
Orotidine	Pyrimidine Metabolism, Orotate containing	-0.4	4.18E-05	3.14E-02
Sphingosine 1-phosphate	Sphingolipid Metabolism	-0.6	2.50E-07	1.88E-04
Sphinganine-1-phosphate	Sphingolipid Metabolism	-0.5	3.17E-05	2.38E-02
4-cholesten-3-one	Sterol	-0.6	1.73E-07	1.30E-04
Alpha-tocopherol	Tocopherol Metabolism	0.5	1.32E-05	9.89E-03
Serotonin	Tryptophan Metabolism	0.6	5.29E-08	3.97E-05
N-formylphenylalanine	Tyrosine Metabolism	-0.5	3.14E-05	2.36E-02
2-oxoarginine	Urea cycle; Arginine and Proline Metabolism	0.7	7.27E-11	5.46E-08
Ornithine	Urea cycle; Arginine and Proline Metabolism	-0.6	6.34E-08	4.76E-05
Proline	Urea cycle; Arginine and Proline Metabolism	-0.5	8.23E-06	6.18E-03
Retinol (Vitamin A)	Vitamin A Metabolism	0.8	1.09E-13	8.21E-11
3,7-dimethylurate	Xanthine Metabolism	0.5	3.28E-05	2.46E-02
3-methylxanthine	Xanthine Metabolism	0.5	5.76E-05	4.33E-02

An OPLS-DA comparing low-moderate versus high CD classes was performed. The statistical model revealed well-explained variance and predictability and the diagnostic performance assessed by leave-one-out analysis exhibited 97.9% sensitivity and 87.8% specificity. The model revealed one class-discriminatory component accounting for 71% of the variation in the data due to increased CD (R -squared- $Y = 0.71$, Q -squared = 0.52) (Fig. 2a). The corresponding loading score, shown in Fig. 2b, identified diacylglycerols, gamma glutamyl amino acids, monohydroxy fatty acids, leucine, isoleucine and valine metabolites, PC and PE as the main metabolites that provide a clear separation between low-medium vs high CD.

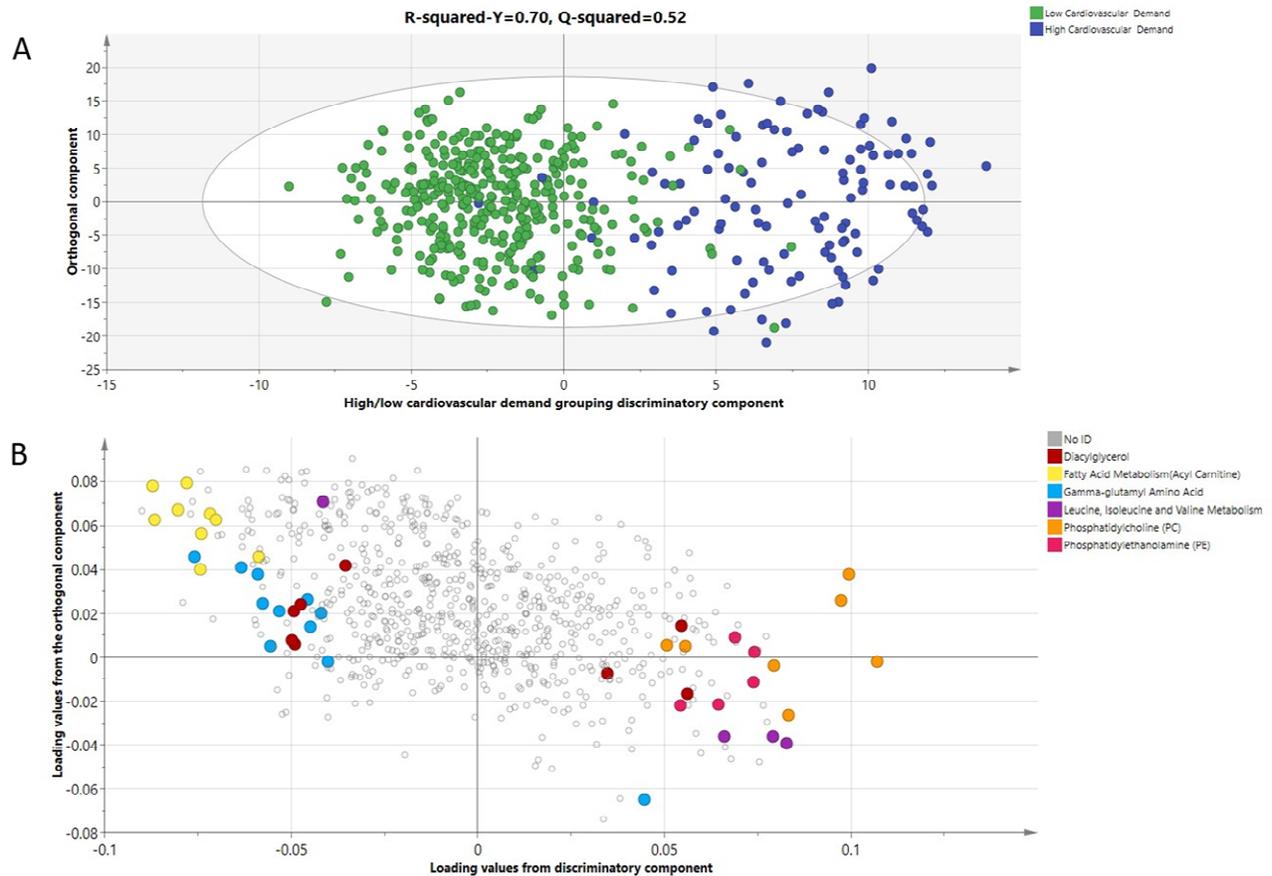


Figure 1. OPLS-DA model comparing low-moderate vs high cardiovascular demand (CD) classes of elite athletes. A. A score plot showing the class-discriminatory component (x-axis) versus orthogonal component (y-axis). B. The corresponding loading plot showing a clustering of PC and PE at the high end of CD opposed by a clustering of diacylglycerols fatty acids (acyl carnitines) and gamma-glutamyl amino acids at the negative end.

GGM networks were constructed using all 751 metabolites, resulting in 60 subnetworks containing >2 metabolites with an overall 600 edges connecting 604 metabolites (nodes) as shown in supplementary Figure 1. The identified subnetworks were filtered for metabolites associated with CD at ($p < 0.05$), resulting in 11 subnetworks containing >2 Bonferroni significant metabolites. Six subnetworks that captured the major metabolic pathways perturbed in relation to CD were selected (Fig 2), including phosphatidyls (cholines and ethanol amines) (Fig 2A), eicosanoids (Fig 2B), Carnitine metabolism (Fig 2C), Gamma glutamyl amino acids and their link to glutamate (Fig 2D), corticol metabolism (Fig 2E), energy metabolites including creatine and TCA (Fig 2F).

Figure 2: Six GGM subnetworks indicate metabolic networks that varied significantly between athletes belonging to low-moderate vs high CD groups. Changes are represented by nodes with sizes proportional to $-\log p$ value (larger nodes indicate more significant association with CD group). phosphatidyls including cholines and ethanol amines (A), ecasinoids (B), carnitine metabolism (C), gamma glutamyl aminoacids and their link to glutamate (D), corticol metabolism (E), energy metabolites including creatine and tricarboxylic acid cycle (F). Red color indicates that high CD has higher metabolite levels, and green indicates that low-moderate CD has higher metabolite levels.

Thirty two metabolites representing different sub-pathways that significantly changed between the two CD groups were projected on the heatmap in Fig 3. The heatmap gives a snapshot summary of the actual intensities of these metabolites after correcting for confounders in the linear model described earlier. Samples were ordered by CD groups into high CD, low-moderate CD 1 and low-moderate CD 2. Whereas the former two groups showed similar intensities of metabolites, the latter group was clearly different, suggesting a presence of a group of athletes with the low-moderate CD group exhibiting a similar metabolic profile of the high CD group than their own low-moderate CD group.

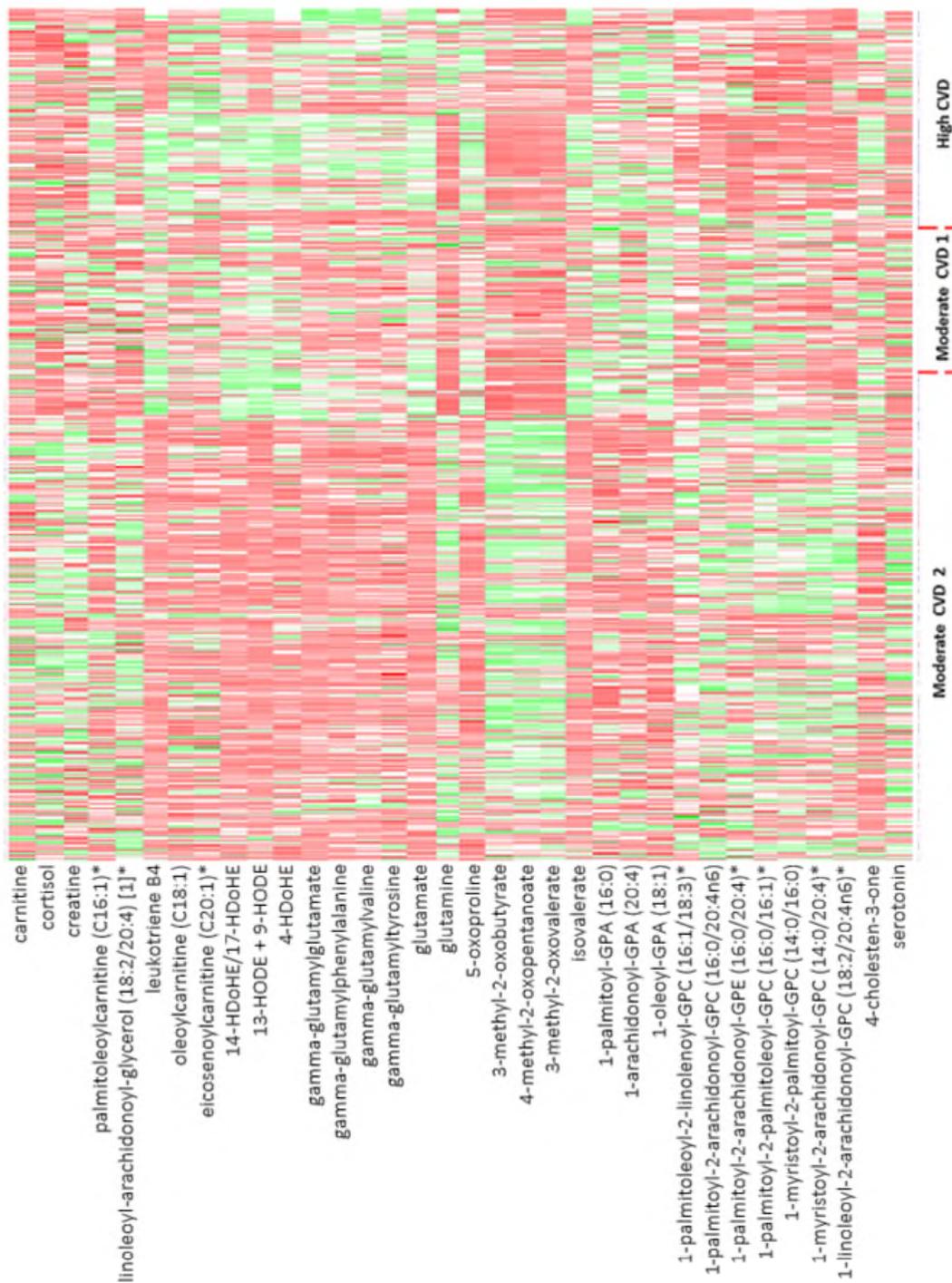


Figure 3. Heatmap of metabolites significantly associated with CD from the linear model association analysis (x-axis). Samples on y-axis were ordered by sports group (high CD, low-moderate CD1, low-moderate CD2). The color code denotes z-scaled values of metabolites after correction of confounders (red represents an increase in high CD, green represents a decrease in high CD).

Discussion

The intensive exercise of elite athletes can cause physiological adaptation of the cardiovascular system leading to increased stroke volume and blood pressure in order to enhance performance. Metabolomics profiling of elite athletes with different CD could provide vital information about their systemic metabolic changes with impact on health and performance [4]. In this study, a non-targeted metabolic profiling of elite athletes who participated in national or international sports events was performed, followed by comparison of metabolic signatures between athletes who belong to a low-moderate CD group versus those who belong to a high CD group. Despite limited information about the participants, the emerging data revealed significant differences in metabolite levels between the two studied groups including differences in energy utilization, oxidative stress scavenging mechanisms, and membrane dynamics. Metabolic changes in high CD group may suggest increased cardiovascular risk, potentially due to exercise-induced left ventricular hypertrophy and increased blood pressure [22].

Metabolites increased in athletes with high CD

Forty metabolites showed higher levels in athletes with high CD compared to their low-moderate CD counterpart. Among the elevated energy-related metabolites in the high CD group was adenine, the building block of ATP used in cellular metabolism as the main source of energy. Creatine was also increased in the high CD group. It plays a critical role in ATP recycling primarily in muscles via donation of phosphate groups to adenosine diphosphate (ADP). Creatine also acts as a pH buffer in skeletal muscle tissues [23], therefore it is expected to be elevated with intensive training associated with increased CD. Interestingly, levels of imidazole lactate were positively correlated with levels of creatine as indicated in the GGM subnetworks, perhaps reflecting increased accumulation of lactate in exercising muscle. Lactate is also used in energy generation through its oxidation to pyruvate by well-oxygenated muscle and heart cells, followed by re-entry into tricarboxylic acid (TCA) cycle [24]. Glutamine was also increased in high CD group. It is mostly synthesized in the muscle tissue, accounting for 90% of all synthesized glutamine, where it can serve as a source of cellular energy next to glucose [25]. Carnitine was also increased in the high CD group. It is mostly accumulated in exercising skeletal muscles where it acts as a transporter of long-chain fatty acids into the mitochondria to be oxidized and produce energy [26].

In addition to energy-related metabolites, specific signaling molecules were increased in high CD group. Among these, three different diacylglycerols containing arachidonic acid (C20:4) were elevated in athletes with high CD including linoleoyl-arachidonoyl-glycerol, palmitoyl-arachidonoyl-glycerol and palmitoleoyl-arachidonoyl-glycerol. Skeletal muscle is an active site of arachidonic acid retention, accounting for 10-20% of the phospholipid fatty acid content on average [27]. In addition to its important function as a second messenger involved in regulating various signaling enzymes (PLC- γ , PLC- δ , and PKC- $\alpha/\beta/\gamma$), arachidonic acid plays a critical role as an inflammatory intermediate [27].

Plasmalogens were too found to be increased in high CD athletes compared to low-moderate counterparts. They represent up to 20% of the total phospholipid mass in humans and $\geq 50\%$ of the ethanol amines fraction in the brain, heart, neutrophils and eosinophils [28]. They play key roles as signaling molecules and modulators of membrane dynamics, providing unique structural

properties, mediating various signaling processes and guarding membrane lipids from oxidation [28]. GGM subnetworks revealed positive correlations among various plasmalogens, GPEs and GPCs, confirming the biochemical relationship among these various phosphatides.

Cortisol was also increased in high CD. It belongs to the glucocorticoid class of hormones released in response to stress and low blood glucose. It plays important roles in stimulating gluconeogenesis to increase blood sugar [29], immunomodulation and metabolism of fat, protein, and carbohydrates [30]. Cortisol is also essential for maintenance of normal blood pressure and in excess can cause hypertension [31]. Therefore, the elevated cortisol in high CD athletes could reflect their response to increased stress and low blood glucose, causing increased blood pressure in this group of athletes. GGM subnetworks reveal a positive correlation between cortisol and other steroids including pregnenolone, an intermediate in the biosynthesis of most of the steroid hormones, suggesting increased steroid biosynthesis shown previously in endurance athletes [19].

An increase in branched chain amino acids (BCAAs, leucine, isoleucine and valine) metabolites was also evident in the high CD group of athletes as manifested by elevated levels of 3-methyl-2-oxovalerate, 3-methyl-2-oxobutyrate and 4-methyl-2-oxopentanoate. Exercise promotes energy expenditure and promotes oxidation of BCAAs that play an important role as substrates to TCA intermediates and gluconeogenesis [32]. Leucine also plays a role as a regulator of intracellular signaling pathway, promoting muscle-protein synthesis in vivo [33].

Metabolites increased in low-moderate CD group of athletes

Seventy metabolites showed higher levels in athletes belonging to the low-moderate CD group compared to their counterparts in the high CD group, including metabolites marking energy generation, oxidative stress and sterol biosynthesis.

Elevation in a number of diacylglycerols, fatty acid-carnitines and acylated carnitines suggest enhanced hydrolysis of diacylglycerols, followed by transfer of fatty acids inside the cells for oxidation and energy production [34]. These changes suggest that low-moderate CD athletes exhibit a higher beta oxidation of fatty acids for energy generation, thus a greater potential to activate lipolysis during exercise than their counterpart in the high CD group. Additionally, accumulation of acylated carnitine may reflect a superior exercise recovery than athletes in the high CD group as carnitine can lower plasma lactate and prolonged exhaustion [35]. Interestingly, four DAGs containing oleic acid, C18:1 (arachidonic acid precursor) were increased in this group (Table 2). The hypotensive effect of oleic acid may partially explain its increase in this group that is characterized by lower blood pressure [36]. Additionally, eicosanoids (products of arachidonic acid) were too increased in athletes with low-moderate CD including 5 & 12 HETE and Leukotriene B4 and B5. Eicosanoids play a crucial role in various physiological and pathological processes such as immunomodulation and platelet aggregation [37, 38].

Other metabolites were higher in low CD group including glutamate, beta-citrylglutamate, gamma glutamyl amino acids and 5-oxoproline, a metabolite of glutathione cycle, suggesting enhanced anti-oxidative stress scavenging mechanism in this group of athletes [39]. Indeed GGM subnetworks showed a strong positive correlation between glutamate gamma glutamylglutamate

and among various gamma glutamyl amino acids, suggesting a biochemical and functional relationship.

Risk of cardiovascular disease

Although exercise is generally known to improve well-being, it can also increase the risk of heart disease including arrhythmias, myocardial infarction, aortic dissection and sudden cardiac arrest [40]. Elevations in several metabolites associated with high CD including BCAAs, PC, and PE may also suggest an increased risk of cardiovascular disease among athletes who belong to this group. Whereas BCAAs were previously correlated with increased risk of cardiovascular disease [41], PC were shown to be associated with increased cardiovascular mortality independent of traditional risk factors [42]. PE (precursors of PC) were too shown previously to be among the strongest predictive lipid species for risk of cardiovascular disease [43]. Whether these associations are truly predictive of increased risk of cardiovascular disease or just a reflection of differences in exercise and dietary requirement in high CD athletes compared to their low-moderate counterpart remains to be investigated. However, it is plausible to assume that high CD athletes could manifest markers of higher risk of cardiovascular disease due to their intensive physical training that is associated with electrical, structural and functional myocardial adaptations [2, 3].

Study limitations

A batch effect related to sample collection from multiple sites was unavoidable. Sources of batch effect include process and time (IN or OUT of competition) of blood collection and duration of transportation and storage, potentially influencing metabolic profiling of collected samples [44, 45]. In spite of these potential confounders, significant metabolic signatures were found following correction for known confounding factors such as hemolysis level and ethnicity. The limited information about participating athletes such as their age, BMI, dietary intake and training regiments has too potentially reduced the power to identify associated metabolites. However the generally young age of elite athletes in addition to their diverse sports disciplines may have reduced the impact of their confounding effects. Additionally, the lack of information related to athletes who belong to team sports with respect to their roles was another limitation. In soccer for example, defenders, mid-fielders and strikers vary in their cardiovascular demand due to their intensity of exercise and style of playing such as change of direction known to elicit greater metabolic and cardiovascular demands compared to straight line running at the same mean speed [46, 47]. This may explain the presence of a low-moderate CD group (moderate CD 1) that shows similar metabolic pattern to the high CD group in Fig 3. This group is mainly composed of soccer players, some of whom potentially exhibit similar exercise patterns and playing style to those who belong to high CD group. Finally, variations in dietary intake between the studied groups such as supplements, medications, and other ergogenics, may have affected their metabolic profile [48]. These differences are challenging to account for as they vary among athletes and are not usually disclosed. Taking all these limitations on board, our data require further confirmation and validation in other sport groups.

Conclusion

Findings of this study present a snapshot summarizing differences between two groups of elite athletes with different cardiovascular demand. These metabolic differences were mostly related to source of energy, mechanisms for scavenging oxidative stress and membrane dynamics. The data also suggest an increased cardiovascular risk in the high CD group compared to their low-moderate counterpart. Replication studies are needed to confirm these metabolic differences in independent data set, aiming for discovery of biomarkers for assessing health, performance, and recovery of elite athletes. Such biomarkers could be used as early signs of extreme trainability related to elite athletic performance with potential applications in guiding future training programs, avoiding potential disorders associated with excessive training as well as improving their overall performance.

References

1. Swann, C., A. Moran, and D. Piggott, *Defining elite athletes: Issues in the study of expert performance in sport psychology*. Psychology of Sport and Exercise, 2015. **16**(1): p. 3-14.
2. Chevalier, L., et al., *Athlete's heart patterns in elite rugby players: effects of training specificities*. Arch Cardiovasc Dis, 2013. **106**(2): p. 72-8.
3. Barbier, J., et al., *Sports-specific features of athlete's heart and their relation to echocardiographic parameters*. Herz, 2006. **31**(6): p. 531-43.
4. Gonzalez-Alonso, J., *Human thermoregulation and the cardiovascular system*. Exp Physiol, 2012. **97**(3): p. 340-6.
5. Rowell, L.B., *Human cardiovascular control*. 1993, New York: Oxford University Press. xv, 500 : ill. ; 24 cm.
6. Laughlin, M.H., *Cardiovascular response to exercise*. Am J Physiol, 1999. **277**(6 Pt 2): p. S244-59.
7. Mitchell, J.H., et al., *Task Force 8: classification of sports*. J Am Coll Cardiol, 2005. **45**(8): p. 1364-7.
8. Fagard, R., *Athlete's heart*. Heart, 2003. **89**(12): p. 1455-61.
9. Zivkovic, A.M. and J.B. German, *Metabolomics for assessment of nutritional status*. Curr Opin Clin Nutr Metab Care, 2009. **12**(5): p. 501-7.
10. Yan, B., et al., *Metabolomic investigation into variation of endogenous metabolites in professional athletes subject to strength-endurance training*. J Appl Physiol (1985), 2009. **106**(2): p. 531-8.
11. Nicholson, J.K., et al., *Metabonomics: a platform for studying drug toxicity and gene function*. Nat Rev Drug Discov, 2002. **1**(2): p. 153-61.
12. Heaney, L.M., K. Deighton, and T. Suzuki, *Non-targeted metabolomics in sport and exercise science*. J Sports Sci, 2017: p. 1-9.
13. Lewis, G.D., et al., *Metabolic signatures of exercise in human plasma*. Sci Transl Med, 2010. **2**(33): p. 33ra37.
14. Goodwin, M.L., et al., *Blood lactate measurements and analysis during exercise: a guide for clinicians*. J Diabetes Sci Technol, 2007. **1**(4): p. 558-69.
15. Berton, R., et al., *Metabolic time-course response after resistance exercise: A metabolomics approach*. J Sports Sci, 2017. **35**(12): p. 1211-1218.
16. Dudzinska, W., et al., *Adenine, guanine and pyridine nucleotides in blood during physical exercise and restitution in healthy subjects*. Eur J Appl Physiol, 2010. **110**(6): p. 1155-62.
17. Howarth, K.R., et al., *Effect of endurance training on muscle TCA cycle metabolism during exercise in humans*. J Appl Physiol (1985), 2004. **97**(2): p. 579-84.
18. Peake, J.M., et al., *Metabolic and hormonal responses to isoenergetic high-intensity interval exercise and continuous moderate-intensity exercise*. Am J Physiol Endocrinol Metab, 2014. **307**(7): p. E539-52.
19. Al-Khelaifi, F., et al., *A pilot study comparing the metabolic profiles of elite-level athletes from different sporting disciplines*. Sports Med Open, 2018. **4**(1): p. 2.
20. Kanehisa, M. and S. Goto, *KEGG: kyoto encyclopedia of genes and genomes*. Nucleic Acids Res, 2000. **28**(1): p. 27-30.
21. Krumsiek, J., et al., *Gaussian graphical modeling reconstructs pathway reactions from high-throughput metabolomics data*. BMC Syst Biol, 2011. **5**: p. 21.
22. Lovic, D., et al., *Left ventricular hypertrophy in athletes and hypertensive patients*. J Clin Hypertens (Greenwich), 2017. **19**(4): p. 413-417.
23. Wallimann, T., M. Tokarska-Schlattner, and U. Schlattner, *The creatine kinase system and pleiotropic effects of creatine*. Amino Acids, 2011. **40**(5): p. 1271-96.

24. Cruz, R.S., et al., *Intracellular shuttle: the lactate aerobic metabolism*. ScientificWorldJournal, 2012. **2012**: p. 420984.
25. Aledo, J.C., *Glutamine breakdown in rapidly dividing cells: waste or investment?* Bioessays, 2004. **26**(7): p. 778-85.
26. Roepstorff, C., et al., *Regulation of plasma long-chain fatty acid oxidation in relation to uptake in human skeletal muscle during exercise*. Am J Physiol Endocrinol Metab, 2004. **287**(4): p. E696-705.
27. Korotkova, M. and I.E. Lundberg, *The skeletal muscle arachidonic acid cascade in health and inflammatory disease*. Nat Rev Rheumatol, 2014. **10**(5): p. 295-303.
28. Braverman, N.E. and A.B. Moser, *Functions of plasmalogen lipids in health and disease*. Biochim Biophys Acta, 2012. **1822**(9): p. 1442-52.
29. Khani, S. and J.A. Tayek, *Cortisol increases gluconeogenesis in humans: its role in the metabolic syndrome*. Clin Sci (Lond), 2001. **101**(6): p. 739-47.
30. Christiansen, J.J., et al., *Effects of cortisol on carbohydrate, lipid, and protein metabolism: studies of acute cortisol withdrawal in adrenocortical failure*. J Clin Endocrinol Metab, 2007. **92**(9): p. 3553-9.
31. Kelly, J.J., et al., *Cortisol and hypertension*. Clin Exp Pharmacol Physiol Suppl, 1998. **25**: p. S51-6.
32. Shimomura, Y., et al., *Exercise promotes BCAA catabolism: effects of BCAA supplementation on skeletal muscle during exercise*. J Nutr, 2004. **134**(6 Suppl): p. 1583S-1587S.
33. Wolfe, R.R., *Branched-chain amino acids and muscle protein synthesis in humans: myth or reality?* J Int Soc Sports Nutr, 2017. **14**: p. 30.
34. Hoppel, C.L., *Carnitine and carnitine palmitoyltransferase in fatty acid oxidation and ketosis*. Fed Proc, 1982. **41**(12): p. 2853-7.
35. Orer, G.E. and N.A. Guzel, *The effects of acute L-carnitine supplementation on endurance performance of athletes*. J Strength Cond Res, 2014. **28**(2): p. 514-9.
36. Teres, S., et al., *Oleic acid content is responsible for the reduction in blood pressure induced by olive oil*. Proc Natl Acad Sci U S A, 2008. **105**(37): p. 13811-6.
37. Lone, A.M. and K. Tasken, *Proinflammatory and immunoregulatory roles of eicosanoids in T cells*. Front Immunol, 2013. **4**: p. 130.
38. Dennis, E.A. and P.C. Norris, *Eicosanoid storm in infection and inflammation*. Nat Rev Immunol, 2015. **15**(8): p. 511-23.
39. Liu, Y., et al., *Emerging regulatory paradigms in glutathione metabolism*. Adv Cancer Res, 2014. **122**: p. 69-101.
40. Lawless, C.E., et al., *Sports and exercise cardiology in the United States: cardiovascular specialists as members of the athlete healthcare team*. J Am Coll Cardiol, 2014. **63**(15): p. 1461-72.
41. Tobias, D.K., et al., *Circulating Branched-Chain Amino Acids and Incident Cardiovascular Disease in a Prospective Cohort of US Women*. Circ Genom Precis Med, 2018. **11**(4): p. e002157.
42. Zheng, Y., et al., *Dietary phosphatidylcholine and risk of all-cause and cardiovascular-specific mortality among US women and men*. Am J Clin Nutr, 2016. **104**(1): p. 173-80.
43. Stegemann, C., et al., *Lipidomics profiling and risk of cardiovascular disease in the prospective population-based Bruneck study*. Circulation, 2014. **129**(18): p. 1821-31.
44. Yin, P., R. Lehmann, and G. Xu, *Effects of pre-analytical processes on blood samples used in metabolomics studies*. Anal Bioanal Chem, 2015. **407**(17): p. 4879-92.
45. Naz, S., et al., *Method validation strategies involved in non-targeted metabolomics*. J Chromatogr A, 2014. **1353**: p. 99-105.
46. Buchheit, M., et al., *Assessing running economy during field running with changes of direction: application to 20 m shuttle runs*. Int J Sports Physiol Perform, 2011. **6**(3): p. 380-95.

47. Buchheit, M. and P.B. Laursen, *High-intensity interval training, solutions to the programming puzzle. Part II: anaerobic energy, neuromuscular load and practical applications*. Sports Med, 2013. **43**(10): p. 927-54.
48. Williams, M.H., *Dietary supplements and sports performance: introduction and vitamins*. J Int Soc Sports Nutr, 2004. **1**: p. 1-6.