



OPEN

Metabolomic response to collegiate football participation: *Pre-* and *Post-*season analysis

Nicole L. Vike^{1,10}, Sumra Bari^{1,10}, Khrystyna Stetsiv¹, Thomas M. Talavage², Eric A. Nauman^{2,3,4,5}, Linda Papa⁶, Semyon Slobounov^{7,11}✉, Hans C. Breiter^{1,8,11} & Marilyn C. Cornelis^{9,11}✉

Contact sports participation has been shown to have both beneficial and detrimental effects on health, however little is known about the metabolic sequelae of these effects. We aimed to identify metabolite alterations across a collegiate American football season. Serum was collected from 23 male collegiate football athletes before the athletic season (*Pre*) and after the last game (*Post*). Samples underwent nontargeted metabolomic profiling and 1131 metabolites were included for univariate, pathway enrichment, and multivariate analyses. Significant metabolites were assessed against head acceleration events (HAEs). 200 metabolites changed from *Pre* to *Post* ($P < 0.05$ and $Q < 0.05$); 160 had known identity and mapped to one of 57 pre-defined biological pathways. There was significant enrichment of metabolites belonging to five pathways ($P < 0.05$): xanthine, fatty acid (acyl choline), medium chain fatty acid, primary bile acid, and glycolysis, gluconeogenesis, and pyruvate metabolism. A set of 12 metabolites was sufficient to discriminate *Pre* from *Post* status, and changes in 64 of the 200 metabolites were also associated with HAEs ($P < 0.05$). In summary, the identified metabolites, and candidate pathways, argue there are metabolic consequences of both physical training and head impacts with football participation. These findings additionally identify a potential set of objective biomarkers of repetitive head injury.

Recent research has pointed to the potentially deleterious consequences of participating in contact sports, such as American football and soccer^{1–3}. These athletes experience impacts to both body and head during play and brain alterations related to head acceleration events (HAEs) have been documented^{3–6}. Despite these consequences, athletic participation and training has many benefits, including improved cardiovascular endurance and psychosocial development⁷. Given this dichotomy, it is imperative to understand the balance between positive and negative health consequences of contact sport participation.

High-throughput metabolite profiling techniques enable comprehensive studies of an individual's metabolic response to certain conditions^{8,9}. Here, metabolite profiling could provide new mechanistic insight into the dichotomy between exercise benefits and HAE-related consequences. Further, knowledge of an objective 'metabolite signature' reflecting potential changes from baseline may also be used to optimize clinical and epidemiological tools for physical assessment in this population of athletes. Previous studies have observed peripheral metabolite changes in contact athletes, but the relationship of these changes to HAEs has not yet been investigated^{18–10}. Additionally, magnetic resonance spectroscopy, a noninvasive neuroimaging technique, has revealed localized neurometabolic changes in contact sport athletes^{11,12}.

¹Warren Wright Adolescent Center Department of Psychiatry and Behavioral Sciences, Feinberg School of Medicine, Northwestern University, Chicago, IL, USA. ²Department of Biomedical Engineering, University of Cincinnati, Cincinnati, OH, USA. ³Weldon School of Biomedical Engineering, Purdue University, West Lafayette, IN, USA. ⁴School of Mechanical Engineering, Purdue University, West Lafayette, IN, USA. ⁵Department of Basic Medical Sciences, Purdue University, West Lafayette, IN, USA. ⁶Department of Emergency Medicine, Orlando Regional Medical Center, Orlando, FL, USA. ⁷Department of Kinesiology, Pennsylvania State University, University Park, PA, USA. ⁸Laboratory of Neuroimaging and Genetics, Department of Psychiatry, Massachusetts General Hospital and Harvard School of Medicine, Boston, MA, USA. ⁹Department of Preventive Medicine, Northwestern University Feinberg School of Medicine, Chicago, IL, USA. ¹⁰These authors contributed equally: Nicole L. Vike and Sumra Bari. ¹¹These authors jointly supervised this work: Hans C. Breiter, Semyon Slobounov and Marilyn C. Cornelis. ✉email: sms18@psu.edu; marilyn.cornelis@northwestern.edu

In the current study, we assessed the metabolomics of collegiate American football athletes before seasonal play (*Pre*) and within one week after the last game of the season (*Post*). Our primary objective was to identify individual metabolite changes in response to athletic participation to gain insight into biological mechanisms by which football participation may impact performance, both positively and negatively. Our secondary objective was to identify metabolite signatures that discriminate periods of competitive play from periods of off-season rest which thus have potential utility in clinical or epidemiological studies of football and health. Finally, we assessed associations *post-hoc* between HAEs and metabolite change between *Pre* and *Post* to provide insight into metabolic change that may be, in part, due to head impacts. Together, these findings may elucidate a set of objective biomarkers that could be used to flag signs of head injury that are not easily detected with traditional behavioral assessments^{13,14}.

Methods

Participants and sample collection. Twenty-three male collegiate American football athletes were recruited for this study as reported previously¹⁵. These athletes were well-seasoned starters, 16 were non-speed linemen, who experience a high number of HAEs^{16,17}. The study was approved by the Pennsylvania State University Institutional Review Board in accordance with the Declaration of Helsinki and written informed consent was obtained from each subject. Demographic information was obtained from each athlete and confirmed by a team physician: age (mean = 21 ± 1 year), race (12 white, 11 African American), years of play experience (YoE; mean = 11 ± 3 years), player position (16 non-speed linemen, 7 speed), and history of diagnosed concussion (HoC; 9 with positive history)¹⁵. None of the athletes received a concussion diagnosis in the 9 months preceding preseason data collection. Five mL of venous blood was collected from each athlete before contact practices (*Pre*) and within one week following the last regular season game (*Post*). Athletes were not required to fast prior to blood collection given they were enrolled in strict nutritional programs. Samples were placed in a serum separator tube, allowed to clot at room temperature, and then centrifuged. Serum was extracted from each tube and pipetted into bar-coded aliquot tubes. Serum samples were stored at -70 °C until they were transported to Metabolon Inc. (Durham, NC, USA) for blinded metabolite analysis. HAEs were collected at each practice session (i.e., between *Pre* and *Post* blood sampling) using BodiTrak's Head Health Network sensor system. For each athlete, HAEs were quantified as the cumulative number of hits exceeding the threshold $Th = 25G$ and $80G$ (e.g., cHAE25G and cHAE80G for each athlete) as well as the average number of hits exceeding $25G$ and $80G$ per practice session with recorded impacts (aHAE25G and aHAE80G)¹². The HAE monitoring protocol and index derivation are detailed in Supplemental Digital Content. Because the majority of head impacts occur during practices and not games¹⁸, and to avoid disrupting game preparation, HAEs were monitored at practice sessions only.

Metabolomics assay, data acquisition and processing. Serum samples were subject to nontargeted metabolomic profiling using UPLC-ESI-MS/MS as previously described^{19,20} and detailed in the Supplemental Digital Content. Mass spectral peaks, retention times, and *m/z* were used to determine the relative quantities of each metabolite. Missing values were imputed with the observed minimum value following normalization and scaling steps. Individual metabolites that contained more than 50% missing values in both *Pre* and *Post* samples were not included for statistical analysis (64 metabolites). The final 1131 metabolites analyzed in the current study are listed in Supplemental Table S1; of those, 209 have not been identified with a known chemical structure and are indicated with prefix "X-" followed by a number (e.g., X-23665).

Statistical analysis. Statistical analyses were performed using R, SAS version 9.2 (SAS Institute Inc, Cary, NC, USA), MetaboAnalyst, or Matlab using log-transformed metabolite values. We first performed standard principal component analysis (PCA) and multilevel PCA to explore the data and identify any outlier samples²¹. For the latter, we generated a data matrix of the within-person variation by subtracting individual metabolite values from the mean metabolite value of *Pre* and *Post*, per participant, per metabolite.

Univariate analysis: individual metabolite changes in response to football participation. Paired t-tests were used to identify metabolites that differed significantly between *Pre* and *Post*. Statistical significance was defined as $P < 0.05$ and FDR (Q value) < 0.05 . Pathway enrichment analysis was performed using MetaboLync (Metabolon Inc., Durham, NC, USA), with all metabolites and their pre-assigned pathways as background and reference pathways, respectively. Analyses were restricted to the 67 pathways containing at least 5 metabolite members. Correction for multiple hypothesis testing in pathway enrichment analysis was performed using an FDR of 5%. We computed both Pearson correlations and pairwise partial correlations to explore the latent relationships of changes in identified metabolites (Δ metabolite, see Supplemental Digital Content). Correlation networks were constructed using Cytoscape²².

Multivariate analysis: sample classification and predictive metabolite-screening. Multilevel partial least squares discriminant analysis (mPLSDA)²³ was performed to examine whether systemic metabolic changes occurred over a single season of football and which metabolites were the most differentiating biomarkers. The prediction error of the mPLSDA model was determined and expressed in terms of number of misclassifications (NMC) and Q^2 by a fivefold cross model validation (CMV) scheme²⁴. To obtain stable class predictions, and stable metabolite selections, the average result of 20 CMVs was calculated. To validate whether the prediction error of the mPLSDA model was not obtained by chance, a comparison was made with the prediction errors from 1000 randomly permuted data sets representing the H_0 -distribution of no-effect. The football season effect was considered statistically significant if the P value obtained from this permutation test was < 0.05 . To select the

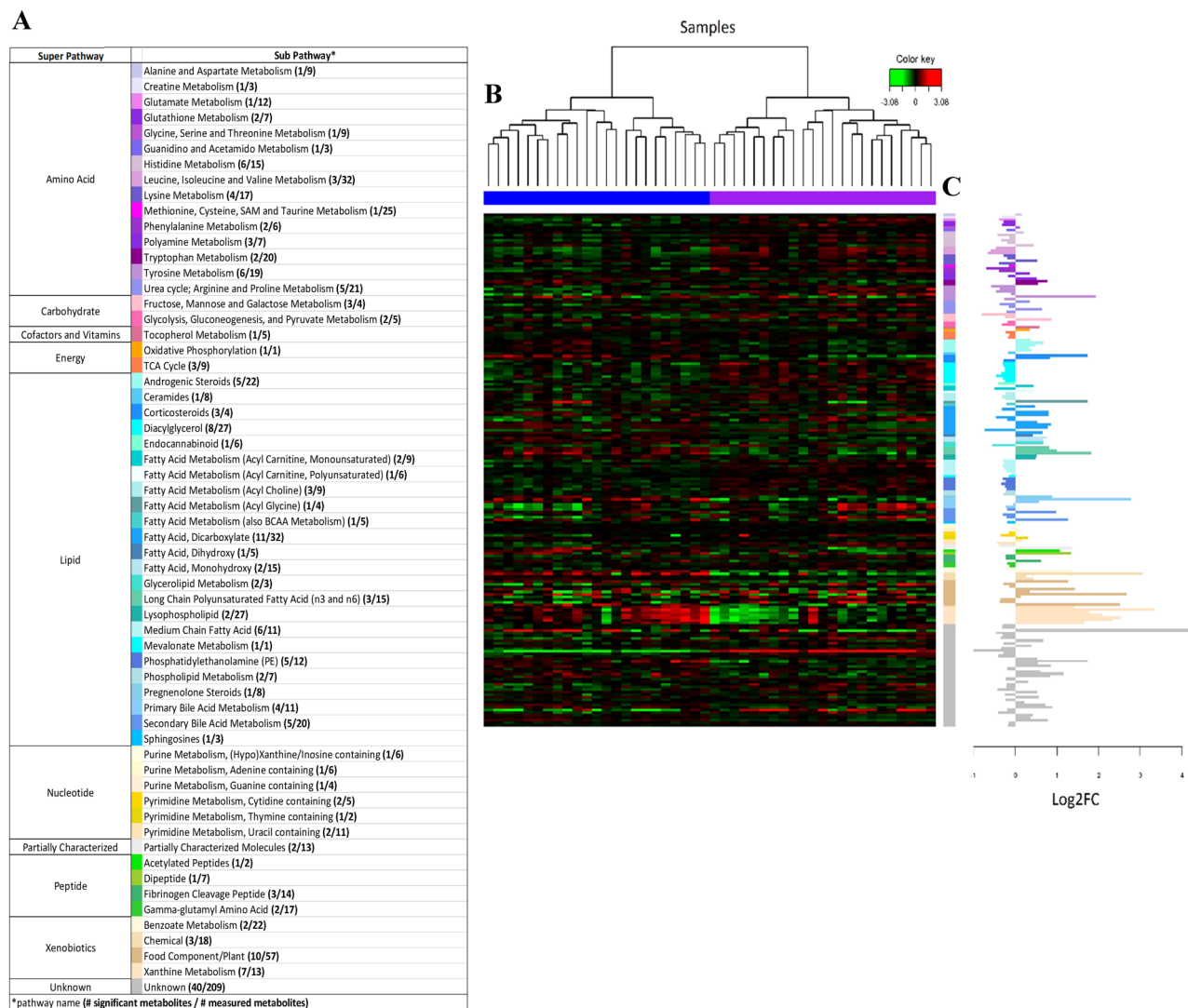


Figure 1. (A) Summary of pathways with metabolites significantly ($P < 0.05$, $Q < 0.05$) changed after a season of football. (B) Multilevel hierarchical clustering (samples only) of the 200 identified metabolites. Green and red cells correspond to low and high metabolite levels, respectively. Columns are samples and rows are metabolites organized by subpathway (see A for color key). (C) Log₂ fold change (FC) for each significant metabolite organized by subpathway (see A) for color key and Supplemental Table S2 for details.

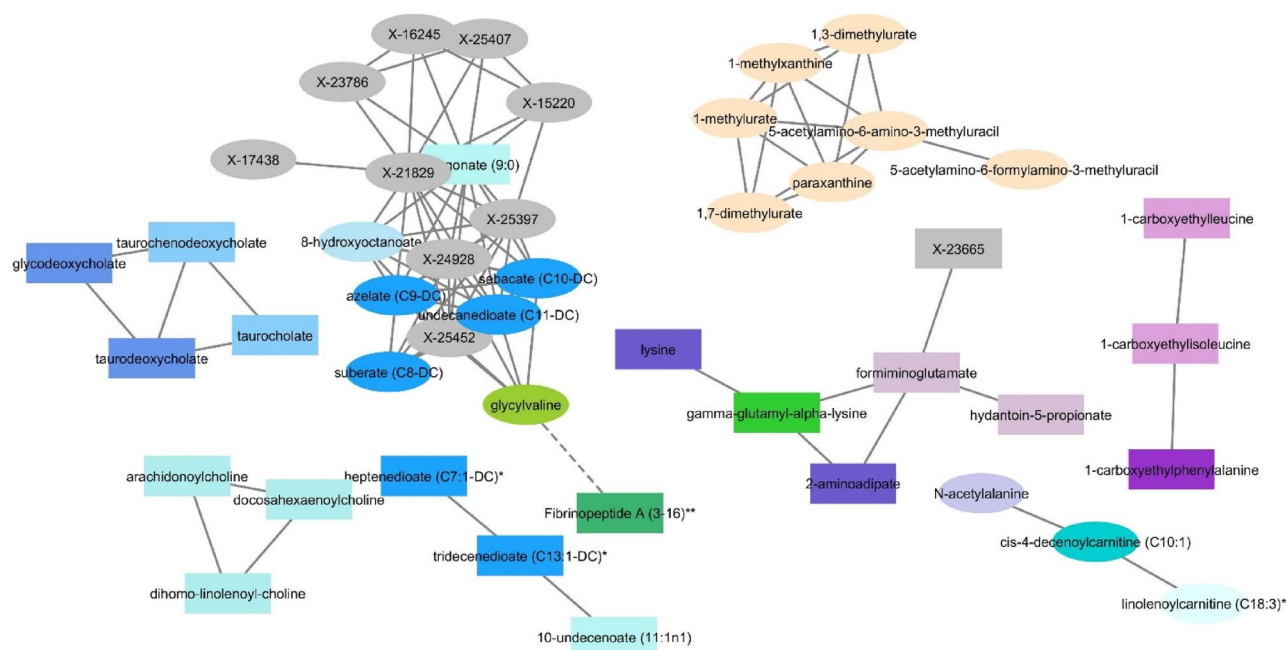
most discriminative metabolites between *Pre* and *Post*, the metabolites were ranked according to their absolute size in the mPLSDA regression coefficient based on the average result of 20 CMVs. Metabolites with the lowest rank product (RP) values have the strongest discriminative power. The RP values of the mPLSDA model were compared with the RP values obtained from 1000 permutations and those with a P value < 0.05 were considered significantly discriminative. Random Forest (RF) analysis, a non-parametric technique unaffected by feature scale, was also implemented as a secondary multivariate analysis (Supplemental Digital Content).

Hierarchical clustering analysis (HCA). Multilevel HCA was used to further visualize pathway enrichment and the discriminatory ability of the significant metabolites identified by univariate and multivariate analysis. HCA was performed using Euclidean distance and complete linkage for grouping of clusters (samples only).

Results

PCA demonstrated separation of *Pre* and *Post* samples along the first component (Supplemental Fig S1). The first component accounted for 9.6% (standard PCA, Supplemental Fig S1a) of the total variance of the data while the multilevel PCA (Supplemental Fig S1b) was able to describe 19.9% of the *within-person* variation. Outliers were vetted for potential technical errors but provided no reason to exclude any samples from our primary analysis.

Individual metabolite changes in response to football participation. A total of 200 metabolites significantly changed over the course of the football season and these mapped to 57 subpathways (Fig. 1, Sup-



Legend

Nodes

Change in metabolite level from *Pre* to *Post*



Edges

— Pearson $r \geq 0.80$
 - - - Pearson $r \leq -0.80$

Color

Super Pathway	Sub Pathway*
Amino Acid	Alanine and Aspartate Metabolism (1/9)
	Histidine Metabolism (6/15)
	Leucine, Isoleucine and Valine Metabolism (3/32)
	Lysine Metabolism (4/17)
	Phenylalanine Metabolism (2/6)
Lipid	Medium Chain Fatty Acid (6/11)
	Fatty Acid, Dicarboxylate (11/32)
	Primary Bile Acid Metabolism (4/11)
	Secondary Bile Acid Metabolism (5/20)
	Fatty Acid, Monohydroxy (2/15)
Peptide	Dipeptide (1/7)
	Fibrinogen Cleavage Peptide (3/14)
Xenobiotic	Gamma-glutamyl Amino Acid (2/17)
Xenobiotic	Xanthine Metabolism (7/13)
Unknown	Unknown (40/209)

*pathway name (# significant metabolites / # measured metabolites)

Figure 2. Pearson correlations (r) of changes in metabolite levels after a season of football. Edges correspond to r and are shown if $|r| > 0.80$. Distances between nodes have no meaning. Only networks of at least 3 members are shown. See Fig. 1A for color key and Supplemental Fig S2 for all networks.

plemental Table S2). Of these metabolites, 90 and 110 demonstrated increases and decreases, respectively. Forty metabolites were of unknown identity.

The metabolomics dataset was significantly enriched for metabolite members of xanthine metabolism ($P = 2.2 \times 10^{-05}$, 3.2-fold-enrichment), fatty acid metabolism (acyl choline, $P = 0.002$, 2.9-fold enrichment), medium chain fatty acid (MCFAs, $P = 0.01$, 2.4-fold-enrichment), primary bile acid metabolism ($P = 0.01$, 2.4-fold enrichment) and glycolysis, gluconeogenesis, and pyruvate metabolism ($P = 0.02$, 3.0-fold enrichment) pathways. Metabolite members of these pathways that achieved only nominal significance (i.e., $P < 0.05$ but $Q > 0.05$) are also listed in Supplemental Table S2.

Focusing only on the 200 significant metabolites, with the exception of three metabolites (1-methyl-4-imidazoleacetate, fructose, X-21467), change in each metabolite was significantly correlated with change in at least one other metabolite based on Pearson $|r| = 0.50$; forming a single network of 197 nodes (metabolites) and 1062 edges (correlations). Applying the more stringent threshold $|r| = 0.80$ resulted in 7 networks of 3 or more nodes that were largely consistent with pathway membership (Fig. 2). Nine metabolites of unknown identity formed a network with pelargonate, 8-hydroxyoctanoate, glycylvaline, and 4 dicarboxylate fatty acids. Supplemental Figure S2 presents the *partial* correlation networks for changes in metabolite levels which, as expected, are much weaker (max $|r_{\text{par}}| = 0.10$) and sparser (77 edges) than their corresponding ordinary correlation networks, because they aim to capture only direct connections between metabolites.

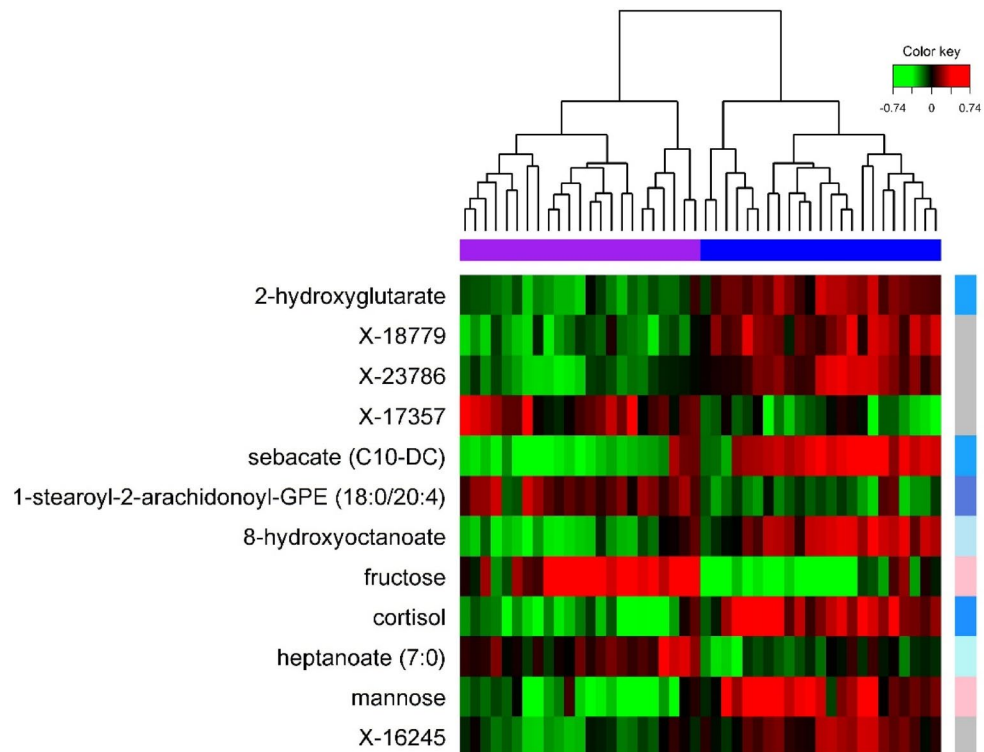


Figure 3. Multilevel hierarchical clustering (samples only) based on the 12 metabolites identified by mPLSDA. Green and red cells correspond to low and high metabolite levels, respectively. Columns are participant samples (i.e., two samples per athlete) and rows are metabolites colored and organized by subpathway (see Fig. 1A for color key).

Pre and Post classification and predictive metabolite-screening. Supplemental Figure S3 presents the results of mPLSDA. On average 2 out of 46 samples (4%) were misclassified ($P < 0.05$ classification model). Significant season effects were also observed based on the Q^2 classification criteria (data not shown). The mPLSDA model yielded 12 metabolites that significantly ($P < 0.05$) discriminated between *Pre* and *Post* (Fig. 3, Supplemental Table S3); all were significant in the univariate analysis. Secondary multilevel RF analysis identified a set of 9 metabolites that yielded optimal classification performance (Supplemental Fig. S4); 5 of these overlapped with the 12 identified by mPLSDA.

Hierarchical clustering analysis. Multilevel HCA based on (1) all 200 identified metabolites (Fig. 1B) or (2) the subset of metabolites showing greatest discriminative performance based on mPLSDA (Fig. 3) demonstrate excellent sample clustering performance.

Post-hoc metabolite and HAE regression analysis. Regression analyses conducted between all 200 identified metabolites and five HAE metrics (number of practice sessions with recorded impacts, $cHAE_{25G}$, $cHAE_{80G}$, $aHAE_{25G}$ and $aHAE_{80G}$) revealed 64 significant ($P < 0.05$) regressions between five HAE metrics and 49 metabolites (Supplemental Table S4). The largest class of metabolites involved in these relationships were lipids, followed by unidentified metabolites, amino acids, xenobiotics, carbohydrates, cofactors, peptides, and energy metabolites. Forty-six of the 64 associations were directionally consistent with those observed for changes *Pre*- and *Post*-season (i.e., metabolites that increased at *Post* also increased with the number of HAE and vice versa).

Discussion

The current study aimed to identify individual metabolite changes in response to collegiate American football participation over a single season. Serum concentrations of 200 metabolites changed and we observed significant enrichment for metabolite members of five pathways. A subset of 12 metabolites was sufficient to accurately discriminate *Post*- from *Pre*-season status and 64 metabolites showed significant associations with HAEs. Our discussion focuses on key metabolic pathways and predictors. Hypotheses connecting other significant metabolites to the *Pre*- and *Post*-season transition are presented in Supplemental Table S2.

Caffeine metabolites (*Xanthine metabolism*) collectively presented with the relatively largest changes *Post*-season. Given known metabolite markers of coffee²⁵ and tea²⁶ consumption did not significantly change, athletes likely increased their caffeine intake through consumption of soda, energy drinks, or supplements. The over five-fold increase in serum saccharin may also suggest some of these sources were sugar-free. To our knowledge, Full Throttle is the only popular sugar-free energy drink containing saccharin in the US; all others are sweetened with

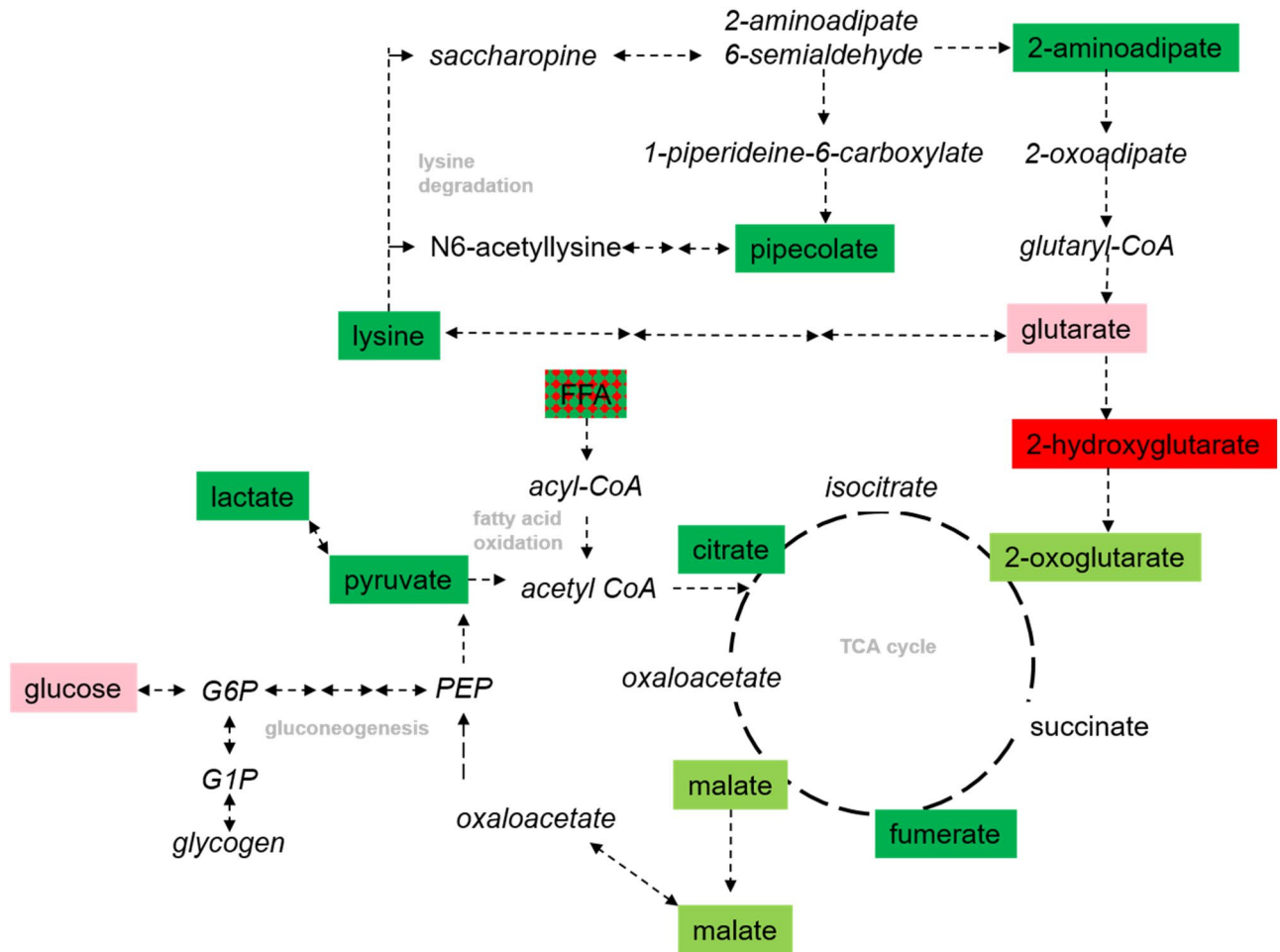


Figure 4. Shifts in energy metabolism from *Pre-* to *Post-*season. Metabolites that significantly ($P < 0.05$, $Q < 0.05$) increased and decreased from *Pre* to *Post* are displayed in red and green boxes, respectively. Corresponding but nominally significant ($P < 0.05$, $Q > 0.05$) metabolites are in light red and light green boxes. Metabolites in *italics* were not measured. Intersecting arrows (\leftrightarrow) imply additional metabolic steps not shown.

acesulfame, aspartame, and/or sucralose²⁷. Caffeine offers a potential edge in athletic performance and is likely the reason for increased caffeine intake by athletes in competition²⁸. Because our athletes were not required to fast prior to blood collection, the marked increase in caffeine metabolite levels may reflect recent or habitual intakes. Accumulating evidence supports a protective role of caffeine against brain injury^{29–31}. Whether habitual caffeine consumption benefits long-term neurological health of football athletes is unknown but warrants consideration.

Most of the assayed *Acyl choline fatty acids* decreased *Post*-season which represent novel findings in the context of chronic physical training and contact sports. Aside from acetylcholine (ACh, not assayed), our knowledge of other acyl cholines is limited. Based on ACh metabolism, we might infer the synthesis of other acyl cholines begins with the activation of the parent fatty acid by an acid-thiol ligase in the presence of coenzyme A and ATP and transfer to choline by choline acetylase³². Arachidonoylcholine, docosahexaenoylcholine, alpha-linoleoylcholine, and oleoylcholine, are reportedly weak to moderate inhibitors of the muscle-type and $\alpha 7$ nicotinic ACh receptor (nAChR)^{33–35}. Since activated $\alpha 7$ nAChR exhibits anti-inflammatory and neuroprotective properties, it is possible that the decreased serum levels of acyl choline fatty acids observed may be a benefit of physical training and exercise, that in turn may offer protection against HAE-related injury during a season of football.

Altered *MCFA* and *Glycolysis*, *gluconeogenesis*, and *pyruvate metabolism* could reflect two etiologies: increased metabolic efficiency incurred with months of training and conditioning over the football season or mitochondrial beta-oxidation dysfunction which would contribute to the decrease in TCA cycle metabolites we observed. Glucose and fatty acids are the dominant fuels oxidized by the muscle for energy production during exercise and the contribution of these fuels can be influenced by diet, muscle glycogen content, exercise intensity, duration, and training status³⁶. Pyruvate, an end product of the glycolysis pathway which feeds into the TCA cycle, lactate, and several TCA cycle intermediates decreased *Post*-season (Fig. 4). Most of the MCFA assayed in the current study decreased, while glycerol and long-chain fatty acids (LCFA) generally increased. These patterns of findings suggest a shift in reliance on glucose to fatty acids as a fuel source which also results in improved ability to maintain glycogen stores^{36,37}. MCFA, in particular, provides a more readily usable energy source than LCFA, possibly due to enhanced cellular uptake and entry into the mitochondria^{38,39}. Indeed, different rates of fatty

acid oxidation might explain why serum MCFAs decreased and LCFAs increased. The former may result from a higher fatty acid oxidation rate than lipolysis rate and vice versa for LCFAs. The increased levels of several oxidized medium and short chain (saturated) fatty acids would support this notion; though the saturation-specificity is more difficult to interpret. In particular, elevated LCFAs may signal mitochondrial beta-oxidation dysfunction which would contribute to the decrease in TCA cycle metabolites we observed. LCFAs (7-hydroxyoctanoate, 8-hydroxyoctanoate, suberate, and sebacate) have been observed to increase in persons with medium chain acyl coenzyme A dehydrogenase deficiency, a rare genetic disorder characterized by dysfunctional mitochondrial beta oxidation^{40,41}.

Primary bile acids are synthesized in the liver, actively secreted into bile and stored in the gallbladder. Bile is discharged into the intestinal lumen in the presence of food and thus bile levels increase after food intake and decrease with fasting⁴². In the current study, cholate and chenodeoxycholate increased while their conjugated metabolites generally decreased. A similar pattern was also observed for several secondary bile acids, which are formed by the intestinal microbiome when some of the primary bile acids enter the colon⁴². Apart from their role as digestive surfactants, bile acids have emerged as important signaling molecules with functions in systemic lipid metabolism, energy metabolism, immune homeostasis, and electrolyte balance⁴². Evidence also suggests that bile acids cross the blood–brain barrier where they might also function in neural health^{43–45}.

Besides changes in the aforementioned metabolic pathways, serum levels of metabolites spanning multiple pathways related to glycerophospholipids also changed *Post*-season. Phosphatidylethanolamines (PEs) levels decreased while others assayed including phosphatidylcholines (PCs), phosphatidylinositols (PI), and phosphatidylserines (PS) did not significantly change. PE is the second most abundant phospholipid in human lipids, after PCs, and higher proportions are found in mitochondria than in other organelles⁴⁶. The majority of PE is synthesized by the CDP-ethanolamine pathway in the endoplasmic reticulum and the PS decarboxylase pathway in mitochondrial inner membranes^{46,47}. Metabolites of the former pathway including phosphoethanolamine, sphinganine-1-phosphate, and several diacylglycerols also decreased. Glycerophospholipids are important structural components of biological membranes⁴⁸. Decreases in PE decrease the fluidity of the membrane and should theoretically compromise membrane integrity and potential⁴⁹. In the mitochondria, PE also facilitates oxidative phosphorylation via the electron transport chain⁴⁶. Endurance exercise training induces proliferation of skeletal muscle mitochondria and recent data suggests a parallel change in mitochondrial phospholipid composition whereby a disproportionate increase in PE occurs compared to other phospholipids⁵⁰. As serum is free of mitochondria, we cannot extrapolate our findings to this organelle and thus the clinical significance of decreased serum PE levels observed in the current study is unknown but warrants further study. Interestingly, the activation of phospholipases that, in turn, degrade cell membranes, is among mechanisms that underlie the neurological damage that accompanies head injury⁵¹. Preliminary human and animal studies report lower plasma levels of PE but also PC and PI with brain injury occurrence¹⁵.

A season of competitive football elicited a plethora of indirect and systemic associations and thus it was not unexpected that a much smaller set of metabolites was sufficient to accurately discriminate periods of competitive play from periods of off-season rest. 2-hydroxyglutarate carried the most discriminative weight and was one of several fatty acid oxidative products that increased *Post*-season. 2-hydroxyglutarate is unique, however, since it is converted to 2-oxoglutarate which then enters the TCA cycle. Since 2-oxoglutarate and several other TCA cycle intermediates at least nominally decreased *Post*-season (Fig. 4), the accumulation of 2-hydroxyglutarate may again reflect greater reliance on fatty acid oxidation for energy. However, this metabolite is also an established oncometabolite that may increase in response to oxidative stress⁵² and thus an alternate mechanism linking this metabolite to football participation cannot be discounted.

Studies exploring exercise-induced alterations of the human metabolome have largely focused on the period up to 24 h after intense and prolonged exercise^{53,54}. Although changes vary by protocol, generally lactate, pyruvate, TCA cycle intermediates, nucleotide degradation products, glycerol, fatty acids, acylcarnitines, and ketone bodies increase after exercise, whereas bile acids decrease. Concentrations of amino acids change in different directions; likely explained by their multiple functions and synthetic pathways⁵⁴. With few exceptions, changes in the metabolome of football athletes after a season of competitive football did not mirror those observed shortly after exercise (Supplemental Table S2) and thus lend confidence to the approach taken to address our study objectives, such as the potential relationship of metabolomic change to contact. Potential metabolite markers of physical fatigue, overtraining, and muscle damage also did not significantly change³⁸. Although cortisol, which increased, is often viewed as having a counter-productive role in exercise due to its catabolic nature, it also plays a key role in exercise adaptations such as stimulation of gluconeogenesis and lipolysis⁵⁵. Urea, a by-product of protein catabolism, did not change. Serum steroid metabolite levels also increased, thus favoring an anabolic over a catabolic state *Post*-season.

As is typical of elite athletes, our sample of football athletes likely maintained some degree of fitness during the off-season and changes in the metabolome reflect football participation, *per se*. Although athlete anthropometry indices were not measured, creatinine, which serves as a rough measure of muscle mass⁵⁶, did not significantly change. Recently, Koay et al.⁵⁷ examined the metabolic effects of an 80-day aerobic and strength exercise intervention in newly enlisted male soldiers. While the sample and study design are somewhat similar to the current study, the intervention did not include physical contact. Few of their findings overlapped with those of the current study (Supplemental Table S2). Although comparable studies are limited, metabolomic patterns reported in the current study are likely unique to athletes engaged in football; a sport characterized by its high contact nature and accumulation of HAEs.

Forty-nine of the 200 metabolites that changed from *Pre* to *Post* also associated with HAE metrics and suggest that the metabolic changes observed might also be related to head impact events that players endure across the football season. The majority of associations involved HAEs exceeding 25G but not 80G; possibly explained

by fewer hits exceeding 80G in this sample. Whether HAEs mediate/moderate *Pre-* and *Post-*season metabolite levels will be the topic of future investigations.

Although the paired-sample design and relatively homogenous sample with respect to age, sex, fitness, and football experience are key strengths of the current study, several weaknesses should be acknowledged in addition to those mentioned above. Given these were elite athletes enrolled in strict nutritional programs, they were not required to fast prior to blood collection. Some of the significant changes in metabolites, non-lipids in particular, may reflect changes in recent food intake prior to *Pre* and *Post* blood collections. Future work will seek to incorporate collection of food frequency diaries to better monitor dietary intake. Given the length of the football season, we cannot rule out an impact of time-varying factors that may induce significant associations. Supervised discriminative techniques tend to over-fit the data and thus our panel of predictive metabolites needs to be validated in an independent sample and further tested for specificity. The current analysis did not account for athlete differences in playing time; this variation will be examined in future analysis as more data become available. In this regard, others have compared metabolomic profiles of athletes sustaining head injuries to comparable healthy athletes, or correlated metabolomic profiles with athlete measures of external load^{8,58,59}. These have highlighted altered blood or urine levels of specific fatty acids, phospholipids, steroids, bile acids, and metabolites of tyrosine and tryptophan metabolism; some of which we observe *Post*-season in the current analysis (see Supplemental Table S2). While the small sample size (N = 23) was adequately powered to detect larger metabolic differences, future work with larger samples may reveal additional, smaller differences.

In summary, our study provides a thorough analysis of the metabolomic changes in response to a full season of collegiate American football participation. While some results support athletes' improved metabolic efficiency, others (i.e., significant metabolite-HAE relationships) suggest potential metabolic changes related to physical contact endured with football participation. The novel metabolites and candidate pathways we have identified may provide new insight to metabolic consequences of the physical training and contact endured with football participation.

Received: 6 October 2021; Accepted: 3 February 2022

Published online: 23 February 2022

References

- Nauman, E. A. *et al.* Post-season neurophysiological deficits assessed by ImPACT and fMRI in athletes competing in American football. *Dev. Neuropsychol.* **40**, 85–91 (2015).
- Breedlove, K. M. *et al.* Detecting neurocognitive and neurophysiological changes as a result of subconcussive blows among high school football athletes. *Athletic Train. Sports Health Care* **6**, 119–127 (2014).
- Bailes, J. E., Petraglia, A. L., Omalu, B. I., Nauman, E. & Talavage, T. Role of subconcussion in repetitive mild traumatic brain injury: A review. *J. Neurosurg.* **119**, 1235–1245 (2013).
- Jang, I. *et al.* Every hit matters: White matter diffusivity changes in high school football athletes are correlated with repetitive head acceleration event exposure. *NeuroImage Clin.* **24**, 101930. <https://doi.org/10.1016/j.nicl.2019.101930> (2019).
- Svaldi, D. O. *et al.* Cerebrovascular reactivity changes in asymptomatic female athletes attributable to high school soccer participation. *Brain Imaging Behav.* **11**, 98–112 (2017).
- Poole, V. N. *et al.* Sub-concussive hit characteristics predict deviant brain metabolism in football athletes. *Dev. Neuropsychol.* **40**, 12–17 (2015).
- Malm, C., Jakobsson, J. & Isaksson, A. Physical activity and sports-real health benefits: A review with insight into the public health of Sweden. *Sports (Basel, Switzerland)* <https://doi.org/10.3390/sports7050127> (2019).
- Daley, M. *et al.* Metabolomics profiling of concussion in adolescent male hockey players: A novel diagnostic method. *Metabolomics* **12**, 185 (2016).
- Fiandaca, M. S. *et al.* Plasma metabolomic biomarkers accurately classify acute mild traumatic brain injury from controls. *PLoS One* **13**, e0195318 (2018).
- Papa, L. *et al.* Elevations in microRNA biomarkers in serum are associated with measures of concussion, neurocognitive function, and Subconcussive trauma over a single national collegiate athletic association division I season in collegiate football players. *J. Neurotrauma* **36**, 1343–1351 (2019).
- Poole, V. N. *et al.* MR spectroscopic evidence of brain injury in the non-diagnosed collision sport athlete. *Dev. Neuropsychol.* **39**, 459–473 (2014).
- Bari, S. *et al.* Dependence on subconcussive impacts of brain metabolism in collision sport athletes: An MR spectroscopic study. *Brain Imaging Behav.* **13**, 735–749 (2019).
- Baugh, C. M. *et al.* Clinical practices in collegiate concussion management. *Am. J. Sports Med.* **44**, 1391–1399. <https://doi.org/10.1177/0363546516635639> (2016).
- Baugh, C. M. *et al.* Concussion management in United States College Sports: Compliance with National Collegiate Athletic Association concussion policy and areas for improvement. *Am. J. Sports Med.* **43**, 47–56. <https://doi.org/10.1177/0363546514553090> (2015).
- Chen, Y. *et al.* Brain perfusion mediates the relationship between miRNA levels and postural control. *Cereb. Cortex Commun.* **1**, tgaa078 (2020).
- Lee, T. A. *et al.* Distribution of head acceleration events varies by position and play type in North American Football. *Clin. J. Sport Med.* <https://doi.org/10.1097/jsm.0000000000000778> (2020).
- Lehman, E. J. Epidemiology of neurodegeneration in American-style professional football players. *Alzheimer's Res. Ther.* **5**, 1–8 (2013).
- Lee, T. *et al.* Head acceleration event metrics in youth contact sports more dependent on sport than level of play. *Proc. Inst. Mech. Eng. Part H J. Eng. Med.* **20**, 0954411920970812 (2020).
- Evans, A. *et al.* High resolution mass spectrometry improves data quantity and quality as compared to unit mass resolution mass spectrometry in high-throughput profiling metabolomics. *Metabolomics* **4**, 2153–0769.1000132 (2014).
- Evans, A., Mitchell, M., Dai, H. & DeHaven, C. Categorizing ion-features in liquid chromatography. *Mass Spectrom. Metab. Data Metab.* **2**, 20 (2012).
- Farnell, D. J., Popat, H. & Richmond, S. Multilevel principal component analysis (mPCA) in shape analysis: A feasibility study in medical and dental imaging. *Comput. Methods Programs Biomed.* **129**, 149–159 (2016).

22. Shannon, P. *et al.* Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504 (2003).
23. Westerhuis, J. A., van Velzen, E. J., Hoefsloot, H. C. & Smilde, A. K. Multivariate paired data analysis: Multilevel PLS-DA versus OPLS-DA. *Metabolomics* **6**, 119–128 (2010).
24. Szymańska, E., Saccenti, E., Smilde, A. K. & Westerhuis, J. A. Double-check: Validation of diagnostic statistics for PLS-DA models in metabolomics studies. *Metabolomics* **8**, 3–16 (2012).
25. Cornelis, M. C. *et al.* Metabolomic response to coffee consumption: Application to a three-stage clinical trial. *J. Intern. Med.* <https://doi.org/10.1111/joim.12737> (2018).
26. van Velzen, E. J. *et al.* Phenotyping tea consumers by nutrkinetic analysis of polyphenolic end-metabolites. *J. Proteome Res.* **8**, 3317–3330 (2009).
27. Heckman, M. A., Sherry, K. & De Mejia, E. G. Energy drinks: An assessment of their market size, consumer demographics, ingredient profile, functionality, and regulations in the United States. *Comp. Rev. Food Sci. Food Saf.* **9**, 303–317. <https://doi.org/10.1111/j.1541-4337.2010.00111.x> (2010).
28. Juliana, M. & Rafaella-Maria, S. Does caffeine enhance athletic performance?. *Arab J. Nutr. Exerc.* **20**, 52–62 (2016).
29. Ning, Y. L. *et al.* Chronic caffeine exposure attenuates blast-induced memory deficit in mice. *Chin. J. Traumatol.* **18**, 204–211. <https://doi.org/10.1016/j.cjtee.2015.10.003> (2015).
30. Lusardi, T. A., Lytle, N. K., Szybala, C. & Boison, D. Caffeine prevents acute mortality after TBI in rats without increased morbidity. *Exp. Neurol.* **234**, 161–168. <https://doi.org/10.1016/j.expneurol.2011.12.026> (2012).
31. Sachse, K. T. *et al.* Increases in cerebrospinal fluid caffeine concentration are associated with favorable outcome after severe traumatic brain injury in humans. *J. Cereb. Blood Flow Metab.* **28**, 395–401. <https://doi.org/10.1038/sj.jcbfm.9600539> (2007).
32. Berry, J. F. & Whittaker, V. P. The acyl-group specificity of choline acetylase. *Biochem. J.* **73**, 447–458. <https://doi.org/10.1042/bj0730447> (1959).
33. Rossier, J. Acetyl-coenzyme A and coenzyme A analogues. Their effects on rat brain choline acetyltransferase. *Biochem. J.* **165**, 321–326. <https://doi.org/10.1042/bj1650321> (1977).
34. Schneider, R. & Timms, A. R. Some aspects of the pharmacology of an homologous series of choline esters of fatty acids. *Br. J. Pharmacol. Chemother.* **12**, 30–38. <https://doi.org/10.1111/j.1476-5381.1957.tb01358.x> (1957).
35. Akimov, M. G. *et al.* Arachidonoylcholine and other unsaturated long-chain acylcholines are endogenous modulators of the acetylcholine signaling system. *Biomolecules* **10**, 283 (2020).
36. Brooks, G. A. & Mercier, J. Balance of carbohydrate and lipid utilization during exercise: The "crossover" concept. *J. Appl. Physiol.* **76**, 2253–2261 (1994).
37. Sidossis, L. S., Wolfe, R. R. & Coggan, A. R. Regulation of fatty acid oxidation in untrained vs trained men during exercise. *Am. J. Physiol.* **274**, E510–515. <https://doi.org/10.1152/ajpendo.1998.274.3.E510> (1998).
38. Lee, E. C. *et al.* Biomarkers in sports and exercise: Tracking health, performance, and recovery in athletes. *J. Strength Cond. Res.* **31**, 2920–2937. <https://doi.org/10.1519/jsc.0000000000002122> (2017).
39. Turner, N. *et al.* Enhancement of muscle mitochondrial oxidative capacity and alterations in insulin action are lipid species dependent: Potent tissue-specific effects of medium-chain fatty acids. *Diabetes* **58**, 2547–2554 (2009).
40. Lee, P. *et al.* L-carnitine and exercise tolerance in medium-chain acyl-coenzyme A dehydrogenase (MCAD) deficiency: A pilot study. *J. Inher. Metab. Dis.* **28**, 141–152 (2005).
41. Bodman, M., Smith, D., Nyhan, W. L. & Naviaux, R. K. Medium-chain acyl coenzyme A dehydrogenase deficiency: Occurrence in an infant and his father. *Arch. Neurol.* **58**, 811–814 (2001).
42. de Boer, J. F., Bloks, V. W., Verkade, E., Heiner-Fokkema, M. R. & Kuipers, F. New insights in the multiple roles of bile acids and their signaling pathways in metabolic control. *Curr. Opin. Lipidol.* **29**, 194–202 (2018).
43. Pan, X. *et al.* Metabolomic profiling of bile acids in clinical and experimental samples of Alzheimer's disease. *Metabolites* **7**, 28 (2017).
44. Quinn, M. *et al.* Bile acids permeabilize the blood brain barrier after bile duct ligation in rats via Rac1-dependent mechanisms. *Dig. Liver Dis.* **46**, 527–534. <https://doi.org/10.1016/j.dld.2014.01.159> (2014).
45. Rodrigues, C. M. *et al.* Neuroprotection by a bile acid in an acute stroke model in the rat. *J. Cereb. Blood Flow Metab.* **22**, 463–471 (2002).
46. Calzada, E., Onguka, O. & Claypool, S. M. Phosphatidylethanolamine metabolism in health and disease. *Int. Rev. Cell Mol. Biol.* **321**, 29–88. <https://doi.org/10.1016/bs.ircmb.2015.10.001> (2016).
47. Schenkel, L. C. & Bakovic, M. Formation and regulation of mitochondrial membranes. *Int. J. Cell Biol.* **20**, 14 (2014).
48. Vance, J. E. & Vance, D. E. *Biochemistry of Lipids, Lipoproteins and Membranes* (Elsevier, 2008).
49. Li, Z. *et al.* The ratio of phosphatidylcholine to phosphatidylethanolamine influences membrane integrity and steatohepatitis. *Cell Metab.* **3**, 321–331. <https://doi.org/10.1016/j.cmet.2006.03.007> (2006).
50. Heden, T. D. *et al.* Mitochondrial PE potentiates respiratory enzymes to amplify skeletal muscle aerobic capacity. *Sci. Adv.* **5**, 8352 (2019).
51. Werner, C. & Engelhard, K. Pathophysiology of traumatic brain injury. *Br. J. Anaesth.* **99**, 4–9 (2007).
52. Yang, M., Soga, T. & Pollard, P. J. Oncometabolites: Linking altered metabolism with cancer. *J. Clin. Invest.* **123**, 3652–3658. <https://doi.org/10.1172/jci67228> (2013).
53. Sakaguchi, C. A., Nieman, D. C., Signini, E. F., Abreu, R. M. & Catai, A. M. Metabolomics-based studies assessing exercise-induced alterations of the human metabolome: A systematic review. *Metabolites* **9**, 164 (2019).
54. Schraner, D., Kastenmüller, G., Schönfelder, M., Römisch-Margl, W. & Wackerhage, H. Metabolite concentration changes in humans after a bout of exercise: A systematic review of exercise metabolomics studies. *Sports Med. Open* **6**, 11. <https://doi.org/10.1186/s40798-020-0238-4> (2020).
55. Hackney, A. C. & Walz, E. A. Hormonal adaptation and the stress of exercise training: The role of glucocorticoids. *Trends Sport Sci.* **20**, 165–171 (2013).
56. Virgili, F. *et al.* Relationship between fat-free mass and urinary excretion of creatinine and 3-methylhistidine in adult humans. *J. Appl. Physiol.* **76**, 1946–1950. <https://doi.org/10.1152/jappl.1994.76.5.1946> (1994).
57. Koay, Y. C. *et al.* Effect of chronic exercise in healthy young male adults: A metabolomic analysis. *Cardiovasc. Res.* <https://doi.org/10.1093/cvr/cvaa051> (2020).
58. Quintas, G. *et al.* Urine metabolomic analysis for monitoring internal load in professional football players. *Metabolomics* **16**, 1–11 (2020).
59. Meier, T. B. *et al.* Prospective study of the effects of sport-related concussion on serum kynurenine pathway metabolites. *Brain Behav. Immunity* **20**, 20 (2020).

Acknowledgements

Data collection was funded by Penn State University. Efforts of NLV, SB, and HCB were funded by the Warren Wright Adolescent Center of Northwestern University. The effort of MCC was supported by National Institute on Aging (K01AG053477). Matlab computations in this paper were run on the Quest cluster supported in part through the computational resources and staff contributions provided for the Quest high performance computing

facility at Northwestern University, which is jointly supported by the Office of the Provost, the Office for Research, and Northwestern University Information Technology.

Author contributions

The initial idea and design for the study was developed, brought through IRB approval, and run by S.S. and H.C.B. S.S. specifically collected the data, with guidance from L.P. for sampling material for omic analysis, and guidance from E.A.N. and T.M.T. regarding collection of head impact data, its quality assurance, and analysis. H.C.B. conceived of the data integration components of this project, contracted vendors for metabolomic analysis and subsequent databasing of the data, and organized the authorship team. M.C.C. and N.L.V. statistically analyzed the derived data and wrote the first draft of the manuscript, with feedback from K.S. and S.B., E.A.N., T.M.T. and H.C.B. All authors critically revised for important intellectual content and approved the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-07079-6>.

Correspondence and requests for materials should be addressed to S.S. or M.C.C.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022