Analytical Methods

ROYAL SOCIETY OF CHEMISTRY

MINIREVIEW

View Article Online
View Journal | View Issue

Cite this: *Anal. Methods*, 2025, **17**, 8847

Evaluating untargeted metabolomics pipelines for sports nutrition research: a review

Yongfu Liu,[ab] Xinghua Liu,[a] Yuting Hu,[a] Feng Gao,[a] Wenjun Yu[a] and Fan Cheng *[a]

Untargeted metabolomics has emerged as a transformative approach in sports nutrition research, offering an unbiased means to characterize the complex biochemical responses to exercise, training, and dietary interventions. Unlike targeted assays restricted to predefined metabolites, untargeted strategies capture broad metabolic perturbations across lipid, carbohydrate, amino acid, and nucleotide pathways, enabling the discovery of novel biomarkers and unanticipated physiological mechanisms. This review critically evaluates the design and application of untargeted metabolomic pipelines in the context of exercise and nutrition science, from pre-analytical sample handling and analytical platforms such as NMR, LC-MS, and GC-MS, to data processing using tools like XCMS, MZmine, and MS-DIAL, and subsequent statistical and bioinformatic interpretation. Key applications include delineating acute metabolic shocks induced by endurance exercise, identifying athlete-specific metabolic phenotypes shaped by chronic training, and assessing the impact of nutritional interventions such as fruit intake, amino acid supplementation, or polyphenol-rich foods on exercise recovery and oxidative stress. The integration of metabolomics with other omics, particularly microbiome metagenomics and lipidomics, highlights the potential for systems-level insights into host–microbe–diet interactions. Nonetheless, significant challenges remain, including the reproducibility of findings, difficulties in metabolite identification, and the translational gap between large datasets and actionable nutritional strategies. By synthesizing current strengths, limitations, and controversies, this review emphasizes that the future of sports metabolomics lies in methodological standardization, multi-omics integration, and validation of candidate biomarkers in independent cohorts. Collectively, these efforts position untargeted metabolomics as a cornerstone for advancing precision nutrition and personalized performance monitoring in athletes.

Received 6th September 2025
Accepted 24th October 2025

DOI: 10.1039/d5ay01484k

rsc.li/methods

Introduction

Sports nutrition research increasingly leverages metabolomics to interrogate the complex biochemical effects of exercise and diet on human physiology.[1] Metabolomics, the comprehensive profiling of small-molecule metabolites in biological samples, provides a snapshot of the metabolic phenotype and how it shifts with acute exercise, training, or nutritional interventions.[2] Unlike traditional targeted assays focusing on a few known biomarkers, untargeted metabolomics is an unbiased, hypothesis-generating approach that casts a wide net over the metabolome.[3] This strategy has revealed unexpected metabolic changes induced by exercise – for example, large post-exercise increases in lipid-derived metabolites and carnitine conjugates – and highlighted how nutritional status modulates these responses.[4] The application of metabolomics in sports has become so prominent that it has given rise to the term "sportomics," denoting the use of omics technologies to study sports and exercise as a model of extreme metabolic stress.[5,6]

However, enthusiasm for untargeted metabolomics is tempered by critical challenges. Detractors argue that untargeted studies produce massive complex datasets with many unknown compounds[7] and can yield irreproducible or hard-to-interpret results.[8] Only a fraction of detected features can typically be confidently identified, due to limitations in spectral libraries and reference standards.[9] Furthermore, subtle differences in sample handling, analytical protocols, and data processing pipelines may lead to divergent findings from the same raw data.[10] Proponents counter that untargeted approaches are indispensable for discovery – for uncovering novel biomarkers and pathways that targeted assays (limited to preconceived metabolites) might miss.[11] They point to successes like the discovery of an exercise-induced metabolic signature resembling the rare disorder hawkinsinuria, or the identification of unique gut microbiome–metabolite interactions in elite athletes, as evidence that untargeted metabolomics can yield meaningful insights into nutrition and performance.[12] The core controversy centers on whether the benefits of broad discovery

[a]*Department of Rehabilitation, Taihe Hospital Affiliated to Hubei University of Medicine, Shiyan, Hubei Province, 442000, China. E-mail: 15071583344@163.com*
[b]*College of Sports Medicine, Wuhan Sports University, Wuhan 430079, China*

outweigh the drawbacks of data complexity and ambiguity in biochemical identification.

In this review, we evaluate untargeted metabolomics pipelines for sports nutrition research. We first outline the typical workflow of untargeted metabolomics – from sample collection to data analysis – and survey the bioinformatic tools available. Next, we examine key findings from recent studies applying metabolomics to exercise and nutrition interventions, highlighting both the insights gained and the recurring patterns observed. We then analyze the major pros and cons debated in the field: what are the strengths of untargeted metabolomics in advancing sports nutrition science, and what are the principal technical and interpretative challenges? The focus of these debates ranges from practical issues (such as data processing variability and metabolite identification bottlenecks) to fundamental questions about how to translate complex metabolomic readouts into actionable nutritional strategies. Fig. 1 provides a high-level view of how metabolomics fits into the broader context of systems biology in exercise, interacting with genomic, transcriptomic, and proteomic factors against the backdrop of environmental influences like training and diet. We aim to provide a balanced perspective and identify where consensus has been reached *versus* where controversies remain. Ultimately, improving and standardizing untargeted metabolomic pipelines – and effectively integrating them with nutritional and physiological data – will be crucial for turning sprawling metabolomic data into genuine advancements in sports nutrition and performance optimization.

Untargeted metabolomics workflow and pipeline tools

Untargeted metabolomics studies generally follow a multi-step analytical pipeline, from careful sample collection to complex data interpretation.[13] Fig. 2 illustrates a typical workflow of an untargeted metabolomic analysis. The process begins with defining the study question and experimental design. Biological samples – commonly blood (plasma/serum) or urine in sports studies, but also muscle biopsies or saliva – are then collected,

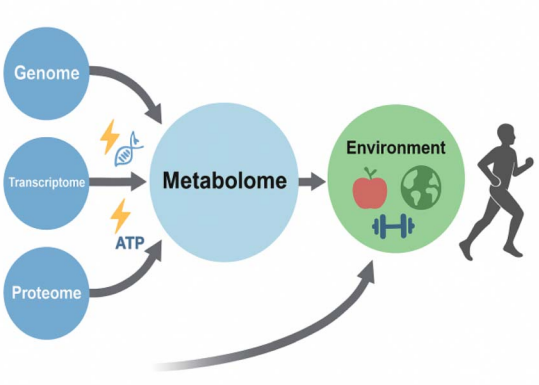

Fig. 1 Metabolomics as an interface between the genome, proteome, and environment in exercise physiology.

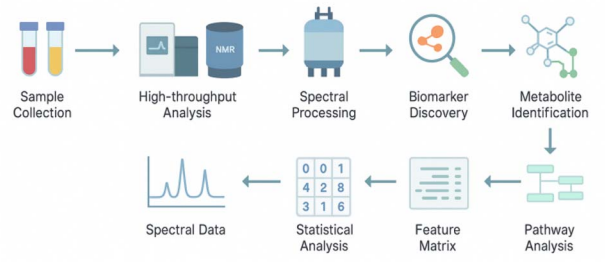

Fig. 2 Analysis workflow in untargeted metabolomic studies. This schematic outlines the key steps of an untargeted metabolomics pipeline, from raw data to biological insight.

with strict protocols to minimize pre-analytical variation.[14] Consistency in timing is especially important: metabolites can fluctuate by 15–40% or more over the diurnal cycle or with recent food intake.[15] For example, delaying post-exercise blood draws or differing fasting status can confound results, as many lipid and amino acid levels exhibit time-of-day or feeding-related swings.[16,17] Attention to such details in the protocol – using the same anticoagulants, standardized clotting times, avoiding hemolysis, and immediate sample freezing – is critical to ensure that observed metabolite changes truly reflect the intervention rather than artefacts. Indeed, a recent review emphasizes that pre-analytical factors can drastically alter metabolomic readouts; for instance, EDTA *vs.* citrate plasma tubes cause significant differences in measured lipid levels. Rigidly standardized sample handling is thus a cornerstone of robust untargeted metabolomics in sports research.

Next, the samples undergo metabolite extraction and analysis using high-throughput platforms, predominantly nuclear magnetic resonance (NMR) spectroscopy or mass spectrometry (MS) coupled to a chromatographic separation.[18,19] NMR-based metabolomics is highly reproducible and non-destructive, providing rich structural information, but has relatively lower sensitivity.[20] MS-based approaches, especially liquid chromatography-MS (LC-MS) and gas chromatography-MS (GC-MS), offer greater sensitivity and coverage of low-abundance metabolites.[21] In practice, many sports metabolomics studies deploy LC-MS or a combination of LC-MS and GC-MS to broaden the range of metabolites detected. For example, among 24 exercise metabolomics papers reviewed by Sakaguchi *et al.*,[22] 11 used LC-MS and several others combined LC-MS with GC-MS, while a few used NMR. The choice of platform can influence the metabolite classes observed – *e.g.* LC-MS excels at lipids, amino acids, and many polar metabolites,[23] GC-MS is strong for organic acids and derivatized compounds, and NMR provides quantitative data on abundant metabolites like lactate, glucose, and certain amino acids.[24] Cutting-edge studies increasingly use multiple platforms or integrate lipidomics (focused MS analyses of complex lipids) alongside global metabolomics to ensure key pathways (like fatty acid metabolism) are comprehensively covered.[25]

In sports-specific matrices and settings, practical choices follow from these strengths. For plasma collected within ~30–90 min after an acute endurance bout (*e.g.*, cardiopulmonary

exercise testing), LC-MS detects transient surges in acylcarnitines and free fatty acids that track mitochondrial β-oxidation dynamics (time-series heatmaps for these classes are reported in a controlled CPX cohort).[26] For urine sampled over the 0–24 h recovery window, GC-MS with methoximation–silylation robustly quantifies TCA intermediates and short-chain organic acids (succinate, citrate, 3-hydroxybutyrate), aiding interpretation of energy metabolism, hydration, and diet effects during recovery. For low-volume, pitch-side saliva and sweat, ^{1}H-NMR and LC-MS are complementary: NMR affords minimal preparation and reliable quantification of abundant metabolites (*e.g.*, lactate, acetate, choline) and has been used to monitor performance in soccer players, while LC-MS—especially with chemical-isotope labeling—expands coverage of amines/phenols in sweat collected after moderate-to-intense exercise and enables practical normalization (*e.g.*, dried-powder mass) for field-collected patches.[27] Finally, during graded exercise, GC-MS/PTR-MS of exhaled breath captures rapid (≈ 3–$5\times$ within $\sim$1 min) rises in VOCs such as isoprene/methyl acetate, providing a non-invasive window into ventilatory/hemodynamic responses under load.[28] Together, these examples operationalize platform selection by matrix and timing: LC-MS for immediate post-exercise plasma lipid/acylcarnitine dynamics; GC-MS for urinary organic-acid profiling during recovery; NMR or LC-MS for on-field saliva/sweat; and GC-MS/PTR-MS for breath-borne VOC kinetics during exercise.

After data acquisition, a crucial and challenging phase is spectral data processing – converting raw spectra into a list of metabolite features and their abundances.[29] For LC-MS data, this typically involves peak detection, deconvolution, and alignment across samples.[30] Commonly used algorithms like XCMS (an R package) detect chromatographic peaks in each sample and align them by *m*/*z* and retention time across the dataset.[31] MZmine 2 is another popular open-source toolkit for LC-MS data processing, offering interactive peak picking, chromatogram building, and alignment.[32] In NMR data, an older approach was simple spectral binning, but this can merge signals; in athlete cohorts, peak-based and deconvolution pipelines are preferred because pre- *vs.* post-exercise pH and ionic changes shift ^{1}H resonances and intensities. Recent sports studies processing urine and plasma by high-throughput ^{1}H-NMR have adopted peak-level quantification and rigorous alignment to resolve exercise-induced changes (*e.g.*, winter-sport athletes around sport-related concussion; recreational athletes before/after HIIT).[33] Regardless of platform, alignment is critical for time-series designs typical in exercise trials: even slight chemical-shift or retention-time drift must be corrected with algorithms such as correlation-optimized or dynamic time warping; we now explicitly point to exercise studies and sports-methods reviews that motivate these steps.[34] Table 1 lists some of the major software tools available for metabolomic spectral processing and data analysis, many of which are free and have been widely adopted in the metabolomics community.

To provide quantitative context, we summarize performance on the widely used MTBLS733 benchmark (Thermo Q Exactive; 836 'true' LC-MS features established by targeted analysis). Using those ground-truth features as the denominator (recall for peak detection), reported recalls were XCMS 820/836 (98.1%), MS-DIAL 799/836 (95.6%), and MZmine 769/836 (92.0%). For quantification accuracy (proportion of detected true features with accurate fold-change), XCMS achieved 89.2% (731/820), MS-DIAL 81.9% (654/799), and MZmine 99.0% (761/769). As a practical precision proxy at the differential-analysis stage, the number of false discriminating markers on this dataset was 51 for XCMS, 42 for MS-DIAL, and 3 for MZmine (true markers: 45, 42, and 48, respectively). These values come from a harmonized re-analysis that directly compared the three tools under matched parameter optimization on MTBLS733.[35] Peak-detection precision/recall has also been evaluated methodologically for XCMS's centWave algorithm on dilution/mixture experiments (seed/leaf extracts), showing higher *F*-scores and improved precision and recall *versus* matchedFilter (XCMS) and centroidPicker (MZmine), consistent with the recall figures above.[36]

Gap-filling and missingness diagnostics are critical for downstream reliability. In a recent modular comparison across common workflows, post-processing missing value fractions were on the order of 2–3% across tools (*e.g.*, $\sim$2.1% for XCMS, $\sim$3.4% for MS-DIAL, $\sim$2.7% for MZmine), underscoring that aggressive gap-filling can reduce missingness but may inflate false positives if not coupled with quality filters. We therefore recommend reporting per-feature gap-filled intensity fractions and excluding features with high gap-filled proportions in athlete-monitoring datasets.[37] For alignment, we now explicitly recommend reporting median ΔRT after warping and the percentage of features with ΔRT below a fixed threshold (*e.g.*, 0.1–0.2 min), alongside the overall missingness after alignment. These diagnostics make the alignment error profile transparent and facilitate comparisons across pipelines in sports plasma/urine studies.[38]

To make QA/QC actionable and comparable across sports-nutrition studies, we adopt the following operational scheme supported by recent guidance and scoping analyses. QC conditioning: inject $\geq$5 pooled-QC samples at the start of each batch to condition the LC-MS system (in demanding matrices, up to 10).[39] In-run pooled-QC frequency: inject one pooled-QC every 6–10 study samples; this reflects current practice (majority of studies) and has minimal impact on coverage while enabling robust monitoring and correction.[40] Acceptance thresholds: (a) feature-level precision: retain only features with QC CV (RSD) $\leq$30% after preprocessing/normalization; remove features exceeding this threshold.[41] (b) Blank exclusion: exclude features with mean process-blank signal $\geq$30% of the pooled-QC signal. (c) Internal-standard performance & system suitability: require IS CV $\leq$ 15% across the run, mass accuracy within $\pm$5 ppm, and retention-time drift $\leq$0.1–0.2 min relative to system-suitability injections; failing IS triggers investigation/re-tuning. Drift correction and diagnostics: apply QC-sample-based correction (*e.g.*, QC-RLSC/LOESS or SERRF) and report quantitative diagnostics before *vs.* after correction: (1) median QC CV% across retained features; (2) PCA of QC samples with the Euclidean distance from the QC centroid (expect tight clustering post-correction); and (3) the proportion of features meeting the CV $\leq$ 30% criterion post-correction. QC-RLSC and

Table 1 Selected open-source software tools for untargeted metabolomics data processing and analysis

Tool	Description & features (analysis step)	Sports-relevant exemplar/context	Primary references
XCMS	LC-MS peak detection, retention-time alignment, feature quantification; supports centroid/profile data; extensive parameterization and CAMERA integration	Used for untargeted plasma LC-MS in acute hypoxic exercise interventions (pre/post/3 h) to characterize exercise-induced metabolic shifts	33
MZmine 2/3	Comprehensive LC-MS workflow (import → detection → deconvolution → alignment → gap-filling → export); modernized in MZmine 3 with multimodal MS support and robust batch processing	Recommended in exercise/sports methods overviews as a robust alternative to XCMS for athlete time-series; MZmine 3 paper documents current best practices	45
MS-DIAL 5	Vendor-neutral GC/LC-MS(/MS) processing with MS/MS-assisted annotation, lipidomics modules, and MSI support; large built-in libraries	Strong fit for exercise lipidomics where training remodels lipid species; also widely used in untargeted workflows relevant to anti-doping contexts	46
OpenMS 3	Modular, reproducible LC-MS(/MS) pipelines (feature finding, alignment, FDR, quantification) with workflow engines (KNIME/Nextflow)	Scalable for longitudinal athlete studies and multi-batch intervention designs requiring reproducible reprocessing	47
MetaboAnalyst 5	End-to-end platform: spectra processing (LC-HRMS), normalization, covariate adjustment, statistics, functional/pathway analysis, and multi-omics integration	Supports covariate handling common in sports trials (*e.g.*, repeated-measures, diet/training covariates) and quick biomarker panel exploration	48
NOREVA 2.0	Web tool to compare normalization strategies with multi-criteria evaluation; supports time-course and multi-class designs	Helpful for pre/post and multi-visit sport designs to select normalization that stabilizes QC CV% and preserves group effects	42
statTarget (QC-RFSC/QC-RLSC)	QC-based signal-drift correction (random forest and LOESS), batch integration, and downstream statistics in R with GUI	Addresses intra/inter-batch drift in large athlete cohorts where frequent QC injections are used	49
SIRIUS + CSI:FingerID (with CANOPUS/COSMIC)	*In silico* formula/structure annotation from MS/MS *via* fragmentation trees and fingerprint prediction; compound class assignment; confidence scoring workflows	Boosts identification levels for exercise-altered unknowns (*e.g.*, acylcarnitines/derivatives) when spectral libraries are sparse	50
GNPS FBMN	Feature-based molecular networking linking LC-MS/MS features by spectral similarity, integrates quantification and ion-mobility; interoperates with XCMS/MZmine/MS-DIAL exports	Useful to visualize families of exercise-responsive metabolites and track isomeric series across training blocks	51
ASICS (R)	Automated ^{1}H-NMR identification/quantification *via* library-based mixture modeling; includes preprocessing and statistical analysis	Appropriate for urine/plasma NMR in athlete monitoring where pH/ionic shifts require deconvolution beyond simple binning	52
BATMAN (R)	Bayesian deconvolution for ^{1}H-NMR spectra accounting for peak shifts; outputs metabolite concentration estimates	Supports high-throughput NMR pipelines in sports cohorts by improving quantification robustness under variable hydration/pH	53

SERRF are established approaches; SERRF has reduced technical error to ~5% RSD in large sets.[42] MS/MS coverage fraction (DDA/DIA acquisition): report the fraction of features (after QC filtering) for which MS/MS spectra were acquired above intensity thresholds, alongside key acquisition settings (*e.g.*, AGC target, NCE, dynamic exclusion). Optimizing DDA parameters materially affects coverage; thus, reporting this fraction increases transparency across studies.[43] Identification-level distribution: report counts (and percentages) of features/metabolites at MSI levels 1–4 (identified → unknown) for the final result set. This provides a transparent view of annotation confidence.[44] Rationale for frequencies and thresholds: the pooled-QC cadence (every 6–10 injections) aligns with a scoping review of current LC-MS untargeted practice; conditioning with multiple QC injections is recommended by recent guidance. A ≤30% QC-CV inclusion rule and ≤30% blank-to-QC ratio are widely used in untargeted workflows, while IS CV ≤ 15%, ±5 ppm mass accuracy and tight RT windows are consistent with system-suitability practice. Together, these criteria provide reproducible, auditable QA/QC suited to sports-nutrition metabolomics.

Once a feature table is generated, data analysis and interpretation can begin. Researchers typically apply a combination of univariate statistics (*e.g.* paired *t*-tests or ANOVA for each metabolite) and multivariate techniques to decipher the metabolomic shifts.[54] In sports metabolomics, multivariate models like PCA and partial least squares-discriminant analysis (PLS-DA) are common to reduce data dimensionality and visualize group separation – for instance, to see if metabolite profiles post-exercise differ from pre-exercise, or if a supplemented group diverges from placebo. Supervised models (PLS-DA, OPLS-DA) have been used to classify athletes *vs.* sedentary controls based on metabolite patterns, or to predict performance metrics from metabolomic signatures. Caution is warranted, as overfitting is a risk with high-dimensional data; best practices include using large sample sizes, cross-validation, and external validation sets to ensure any putative biomarkers are robust. To operationalize this, PLS/OPLS models should be evaluated with repeated or k-fold cross-validation reporting both R^2 and Q^2, complemented by response-permutation testing (*e.g.*, 500–2000 permutations) and CV-ANOVA to test model significance (typical criterion $p < 0.05$). These procedures help detect spurious class separation and optimism in predictive performance; seminal evaluations in metabolomics demonstrate that permutation and CV-ANOVA effectively flag overfitting and that naïve validation yields overly optimistic results.[55] Y-Scrambling (randomly re-assigning class labels and refitting) should further be reported to demonstrate that observed discrimination exceeds that expected by chance.[56]

Because sports studies frequently employ cross-over or longitudinal designs with repeated measures, analysts should either (i) adopt multilevel decompositions that remove between-subject variation prior to PCA/PLS-DA (multilevel sPLS-DA; implemented in mixOmics *via* the within-subject variation), or (ii) perform metabolite-wise linear mixed-effects modelling with random intercepts/slopes for participants, followed by multiple-testing control and, where appropriate, multivariate aggregation of significant features. These approaches are explicitly designed for paired/repeated data and have been shown to improve feature selection and classification by accounting for subject-specific baselines—an important consideration in athlete cohorts.[57] In sport-science more broadly, methodological guidance recommends mixed models for longitudinal, imbalanced datasets typical of training studies; this translates directly to metabolomics where within-athlete correlation structures otherwise inflate false positives.[58]

A final and formidable step is metabolite identification – figuring out what compounds the significant spectral features correspond to.[59] This often requires matching MS/MS fragmentation spectra to databases (like HMDB, METLIN, or MoNA) or confirming with authentic standards.[60–62] In untargeted studies, typically only a subset of features (sometimes ~20–50%) can be confidently annotated.[50] For example, in an untargeted LC-MS study on cyclists, ~509 features were detected but 107 metabolites were "chemically identified" as known compounds.[63] The rest remained putative or unknown. The reliance on database matching means that novel or unexpected metabolites may go unidentified – a key limitation of untargeted metabolomics in general. Some advanced techniques like tandem MS (MS/MS) networking, retention time prediction, or ^{13}C-labeling can help illuminate unknowns, but these add complexity and are not yet routine in sports studies. Scholars have highlighted metabolite identification as a "bottleneck" and called for community efforts in sharing reference spectra

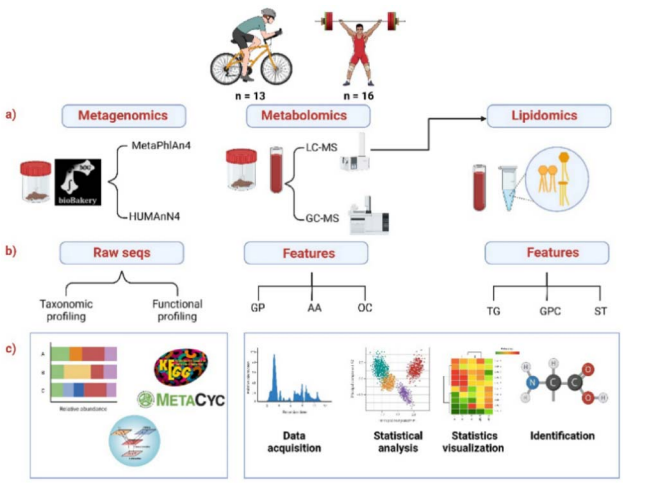

Fig. 3 Multi-omics integration in sports metabolomics. This diagram illustrates an end-to-end pipeline combining metagenomics and metabolomics, as applied in a recent study of elite weightlifters *vs.* endurance cyclists.[67] (a) Samples are collected and subjected to different analytical techniques: fecal samples for metagenomic sequencing of the gut microbiota; blood (plasma) samples for untargeted metabolomics and lipidomics profiling. (b) Data processing converts raw data into features: DNA sequences are processed to identify microbial taxa and genes (using bioinformatics tools like MetaPhlAn), while chromatographic MS data are processed to extract metabolite and lipid features. (c) Data integration and analysis bring together the metagenomic and metabolomic layers. For example, multivariate models (PCA/OPLS-DA) can distinguish athlete groups based on combined omics profiles, and network analysis can link specific microbes with metabolites.

and developing better computational identification tools.[64] In nutrition studies, where many metabolites may derive from diet or gut microbiota, identification is particularly challenging because not all food-derived compounds are well-represented in standard libraries.

Despite these hurdles, untargeted pipelines have successfully catalogued a broad array of metabolic changes with exercise and diet.[65,66] Recent efforts to integrate metabolomics with other "omics" (transcriptomics, proteomics, and especially microbiome metagenomics) are creating a more holistic systems biology approach to sports nutrition. Fig. 3 shows an example of such an integrative workflow from a 2025 study, where fecal metagenomics was combined with plasma metabolomics/lipidomics in elite athletes.[67] By merging multi-omics data, researchers can begin to connect specific gut microbial pathways with circulating metabolite profiles (for instance, seeing how microbial metabolism of amino acids might link to host energy metabolites in endurance *vs.* strength athletes). These comprehensive pipelines underscore that metabolomics is not a stand-alone technique, but one piece of a larger puzzle. In summary, the untargeted metabolomics workflow – from sample prep to bioinformatics and biological interpretation – is complex, but when executed rigorously it provides an unparalleled window into the metabolic adaptations underlying sports performance and nutrition.

Metabolomics applications in exercise and sports nutrition

Untargeted metabolomics has been applied to a wide spectrum of questions in exercise science and sports nutrition, from characterizing the acute metabolic shock of intense exercise to monitoring long-term training adaptations and dietary intervention effects.[68] A consistent finding across many studies is that exercise – especially prolonged endurance exercise – triggers large-scale shifts in the metabolome, predominantly affecting pathways of energy metabolism, lipid oxidation, and redox balance.[69] These metabolic perturbations often mirror known physiology (*e.g.* mobilization of fat stores, activation of glycolysis) but metabolomics allows their quantification and the discovery of novel biomarkers of exercise stress or recovery.[70] On the nutrition side, metabolomics has been used to evaluate how specific foods or supplements (such as polyphenol-rich beverages, amino acid supplements, or dietary patterns) modulate the metabolomic response to exercise or improve metabolic health in athletes.[71] This section reviews major findings under three contexts: acute exercise effects, chronic training and athlete status, and nutritional interventions. Table 2 summarizes some hallmark metabolite changes observed in these scenarios, illustrating the consistency (and variability) of results across studies.

Acute exercise metabolomics

When an individual engages in vigorous exercise, the immediate metabolic response is profound.[72] Untargeted metabolomics studies have repeatedly shown that high-intensity or long-duration exercise causes a surge in circulating lipid-related metabolites, reflecting enhanced lipolysis and fatty acid oxidation to meet energy demands. For example, running a marathon or cycling for 75 km can increase plasma levels of free fatty acids,[73] acylcarnitines (fatty acid oxidation intermediates),[74] and ketone bodies several-fold within an hour or two.[75] Nieman *et al.*[63] reported that after a 75 km cycling trial (water-only), 107 metabolites rose >2-fold (35 of them >5-fold) post-exercise, all of which were lipid-derived species such as long-chain fatty acids, dicarboxylic fatty acids, acylcarnitines, and ketones. By 21 hours of recovery, most of these had returned near baseline, indicating the transient nature of the extreme metabolic shifts. Similarly, other studies of endurance exercise (2+ hours at ~70% VO_2max) found broad increases in fatty acid oxidation products (*e.g.* 3-hydroxybutyrate, octanoylcarnitine) accompanying decreases in triglycerides, as muscle and liver actively consumed blood lipids for fuel. Table 2 shows representative changes: one study noted 88% of measured free fatty acids and 91% of acylcarnitines increased after a 51 km cross-country ski march, whereas 88% of mono- and diacylglycerols decreased, consistent with triacylglycerol breakdown for energy.

Besides lipids, carbohydrate metabolism exhibits expected shifts: intense exercise raises glycolytic intermediates (*e.g.* glucose-6-phosphate, pyruvate)[76] and end-products like lactate,[77] while depleting some TCA cycle intermediates and related compounds as they are rapidly cycled. In moderate 30–60 min exercises, metabolomics has detected significant lactate elevations (confirming anaerobic glycolysis) and increases in Krebs cycle metabolites such as succinate and malate.[78] However, the changes in TCA cycle metabolites are often relatively modest (on the order of <2-fold) compared to lipids. This may be because these intermediates are tightly regulated and also rapidly reused, or due to their compartmentalization in muscle *vs.* blood. Still, even short bouts can perturb them: one study of 18 min high-intensity running found small but detectable increases in plasma succinate and citrate immediately post-exercise. Another recurring theme is amino acid metabolism: exercise tends to decrease plasma levels of certain amino acids (*e.g.* tryptophan, branched-chain amino acids) and increase others or their catabolites (like alanine, glutamine) as part of the alanine-glucose cycle and protein breakdown for energy. Tryptophan in particular often drops during prolonged exercise, likely due to increased uptake and oxidation or its conversion to serotonin and kynurenine; in a cycling fatigue test, plasma tryptophan fell significantly as fatigue approached. Correspondingly, metabolites of purine nucleotide breakdown (indicative of ATP turnover and fatigue), such as hypoxanthine, may rise during high-intensity efforts – one study in runners showed hypoxanthine spiking after exhaustive exercise, reflecting muscle ATP degradation and perhaps contributing to oxidative stress upon its conversion to uric acid.[79]

As seen above, untargeted metabolomics confirms many expected exercise-induced changes but also provides a more nuanced picture.[84] For instance, it was through metabolomic analyses that researchers noted a consistent post-exercise increase in unusual dicarboxylic fatty acids (*e.g.* azelate, sebacate), suggesting amplified ω-oxidation of fatty acids during intense exercise beyond traditional β-oxidation.[85] Metabolomics

Table 2 Characteristic metabolite changes detected by untargeted metabolomics in exercise and sports nutrition studies

Metabolite or pathway	Acute exercise response (single session)	Training or athlete baseline	Nutritional modulation
Free fatty acids (FFA)	↑ FFA during prolonged moderate/high-intensity exercise (2–10 fold post-exercise) as adipose lipolysis accelerates. Example: palmitate, oleate peak 1 h post-run, returning to baseline by 24 h	Endurance-trained athletes often show lower fasting FFA but more rapid FFA mobilization when exercise starts	High-carb meals before exercise attenuate FFA rise (insulin effect), whereas fasted exercise exaggerates FFA release. Fat-rich diets can elevate baseline FFA availability
Acylcarnitines & ketones	↑ Acylcarnitines (medium- and long-chain) and ketone bodies (β-hydroxybutyrate, acetoacetate) after endurance exercise. Indicates enhanced β-oxidation and overflow of acetyl-CoA. In a 75 km cycling, acylcarnitine levels spiked ∼4–6× at 1.5 h post-exercise. Ketones like 3-hydroxybutyrate rose >5×, reflecting hepatic ketogenesis	Trained individuals may have higher peak ketone responses to exercise (due to metabolic flexibility) but also clear them faster	Ketogenic or low-carb diets significantly elevate exercise-induced ketones and acylcarnitines at rest and during exercise. Carbohydrate ingestion blunts ketone formation during exercise
Triglycerides (blood)	↓ Glycerides (triacylglycerols) immediately post prolonged exercise, as they are consumed for fuel. In a multi-hour ski march, ∼88% of detectable di- and triacylglycerols dropped in plasma. They tend to normalize or overshoot during recovery as lipoproteins redistribute lipids	Athletes (especially endurance) generally have lower resting triglycerides than sedentary peers, reflecting training-induced improvements in lipid clearance	Certain supplements (*e.g.* omega-3 fatty acids) can lower resting triglycerides. High glycemic carbs post-exercise can accelerate triglyceride restoration by promoting lipid uptake into tissues
Glucose & lactate	↑ Lactate (often 3–10×) with high-intensity exercise, concurrent with transient ↑ glucose (mobilization from liver glycogen). Example: after 30 min at 110% VO_2max intervals, blood lactate was substantially elevated *versus* moderate exercise. Lactate usually clears within 1–2 h post-exercise. Glucose may dip below baseline during recovery if glycogen depleted	Trained athletes show a smaller rise in blood lactate at a given workload (higher lactate threshold). They also maintain tighter glycemic control during exercise, relying more on fat oxidation	Pre-exercise carbohydrate loading elevates starting glucose and can attenuate early lactate accumulation. Ingredients like caffeine can increase peak lactate by allowing higher exercise intensity for the same effort
Amino acids (BCAA, Trp)	↓ Certain amino acids (*e.g.* branched-chain amino acids valine, leucine, and tryptophan) during exhaustive exercise. Believed to be taken up by muscle for oxidation or converted (tryptophan → serotonin/kynurenine). ↑ Alanine and glutamine often observed (glucose–alanine cycle, ammonia scavenging). Ammonia-related metabolites (glutamine, urea) may rise indicating amino acid catabolism	At rest, endurance athletes sometimes have lower BCAA levels than sedentary (possibly due to chronic utilization or dietary differences). Training can increase the efficiency of amino acid turnover, reflected in faster post-exercise recovery of amino acids	Protein or BCAA supplementation before exercise can increase baseline and post-exercise BCAA levels and may reduce their drop during exercise. Certain supplements (*e.g.* β-alanine) elevate specific amino acid-derived metabolites (carnosine) in muscle, detectable indirectly *via* histidine/alanine changes
TCA cycle & energetics	↑ TCA intermediates (citrate, succinate, fumarate) modestly after exercise, but pattern is complex. Succinate often accumulates (incomplete oxidation in oxygen-limited muscle), then is cleared. Markers of nucleotide breakdown (inosine, hypoxanthine) ↑ with intense exercise, indicating ATP turnover and contributing to oxidative stress post-exercise	Resting levels of TCA intermediates don't differ greatly by training status, but trained muscle may have higher mitochondrial enzyme activity – metabolomics of muscle tissue can show higher baseline TCA cycle intermediate content in athletes. Chronic high-intensity training can lower resting purine breakdown products due to improved energy efficiency	Antioxidant-rich diets might modulate exercise-induced purine oxidation and reduce accumulation of urate/xanthine. Some evidence suggests β-alanine supplementation (to boost carnosine) might indirectly stabilize TCA cycle metabolism by buffering muscle pH, though metabolomic signatures are still being studied

Table 2 (*Contd.*)

Metabolite or pathway	Acute exercise response (single session)	Training or athlete baseline	Nutritional modulation
References	High-intensity/long-duration exercise lipid responses;[80] Ski march FFA and acylcarnitine changes;[81] lactate and carb metabolism;[78] amino acid shifts and ammonia;[82] purine metabolism and fatigue;[6] training *vs.* sedentary metabolite differences[83]		

also draws attention to metabolites in the interface of multiple pathways:[86] an example is 2-hydroxybutyrate, a byproduct of both amino acid catabolism and glutathione synthesis. Several studies observed a decrease in 2-hydroxybutyrate with exercise despite increased lipid use – potentially linked to its consumption in counteracting oxidative stress *via* glutathione.[78,87] Such insights exemplify how untargeted data can generate new hypotheses about exercise metabolism (in this case, that antioxidant demands rise and consume certain intermediates during exercise).

Chronic training and athlete metabolomics

Beyond acute bouts, metabolomics has been applied to understand how long-term training or athletic status alters the metabolome. Cross-sectional studies comparing elite athletes to non-athletes find that habitual exercise is associated with a distinctive metabolic phenotype. Generally, endurance-trained athletes tend to have lower basal levels of metabolites linked to cardiovascular risk (*e.g.* lower triglycerides and LDL-associated lipids, lower branched-chain amino acids), and higher levels of certain beneficial metabolites (like taurine or HDL-associated lipids). For example, Kujala *et al.*[83] showed that physically active adults had more favorable metabolite profiles – including lower unsaturated fatty acids and triglycerides – compared to sedentary controls. Untargeted profiles of team-sport athletes (*e.g.* soccer players) reveal enhanced markers of oxidative metabolism and differences in steroid hormone metabolites reflecting training loads. A notable large study by Al-Khelaifi *et al.*[88] profiled 191 elite athletes from power and endurance sports. They found some commonalities – for instance, athletes across the board showed elevated indicators of oxidative stress and antioxidant activity (signs of repeated exercise-induced stress). However, there were also distinct signatures: endurance athletes had greater increases in metabolites from steroid hormone and polyamine pathways (perhaps related to endurance training stress and recovery), whereas power athletes (sprinters, weightlifters) showed higher levels of certain sterols, purine derivatives, and energy metabolites at rest. This suggests that the metabolic demands of different sports (aerobic *vs.* anaerobic) lead to divergent chronic adaptations detectable in the metabolome. Another investigation of professional soccer players reported that even within a homogeneous group, metabolomics could distinguish those with different training statuses or diets, though intra-group differences were subtle compared to athlete-vs-nonathlete differences.

Intervention studies on training provide dynamic information: for instance, a 6 weeks training study in cyclists (comparing two training intensity distributions) used ^{1}H-NMR metabolomics on urine and found changes in hippurate, creatinine, and other metabolites reflecting shifts in gut-microbiome co-metabolism and energy metabolism post-training.[89] Hippurate (a gut microbial cometabolite of polyphenols) decreased with certain training regimens, suggesting training might influence gut metabolism or dietary intake patterns. While the interpretations can be complex, these studies underscore metabolomics' potential to serve as a comprehensive monitoring tool for athlete conditioning, recovery status, or even overtraining – though more research is needed to establish reliable biomarkers.[90] An emerging concept in personalized nutrition and exercise science is the athlete 'metabotype'—the stratification of individuals into metabolically homogeneous subgroups based on their metabolome (and related clinical/omic data) at baseline or in response to a standardized stimulus. This concept builds on nutrition science, where metabotyping has been proposed and tested as a framework for tailoring diet at the group level and for predicting inter-individual responses to interventions.[91,92] In sports cohorts, metabolic phenotyping has differentiated athlete groups and linked metabolite patterns to functional characteristics: large cross-sectional work in elite and sub-elite athletes shows discipline- and training-status-specific metabolic profiles,[93] targeted panels separate endurance *vs.* strength-trained athletes across a training year,[94] and blood metabolic phenotypes measured around a standardized exercise test discriminate sprint performance tiers, lactate responses, and subsequent illness susceptibility in highly-trained skiers.[95] Moreover, specific metabolite panels correlate with aerobic capacity and race performance, supporting the construct validity of an 'athlete metabotype' (*e.g.*, VO_2max-linked metabolites in runners and differential acylcarnitine/arginine-related signatures across performance strata) [Shi *et al.*, 2020; Schader *et al.*, 2020; Kelly *et al.*, 2020].[96–98] Together, these data justify the use of 'metabotype' in an athlete context while highlighting that rigorous, standardized phenotyping and validation are required before routine deployment. As large datasets are analyzed, we

expect clearer patterns to emerge linking specific metabolomic signatures to fitness phenotypes.

Nutritional interventions and metabolomics

Untargeted metabolomics has also been applied to dietary studies in athletic contexts to evaluate how nutrition influences the metabolome and performance recovery. One approach is to feed different nutrients or supplements and observe the downstream metabolite changes.[99] A striking example is the study by Nieman *et al.*[80] which compared banana ingestion *vs.* pear *vs.* water during a 75 km cycling trial. All trials showed the large exercise-induced lipid metabolite surge, but the fruit-fed trials showed attenuated changes in some markers of inflammation and oxidative stress. Metabolomics revealed subtle differences: for instance, bananas led to higher levels of catechol derivatives post-exercise, whereas water-only led to greater elevations in free fatty acids. These untargeted findings supported the idea that fruit provides not just glucose but also secondary metabolites that can modulate metabolic and inflammatory responses. Another study investigated pistachio ingestion before cycling.[63] While performance did not change, metabolomics showed that pistachio (rich in antioxidants and L-arginine) caused higher plasma arginine and certain amino acid derivatives, and a blunted increase in some lipid oxidation products, compared to control. Fig. 4 shows one outcome from that trial: the plasma level of 9,10-dihydroxy-12-octadecenoic acid (9,10-DiHOME, a linoleate oxidation product) was significantly lower when cyclists ate pistachios, suggesting reduced oxidative stress or altered lipid metabolism. Such markers (DiHOMEs) have been proposed as inflammatory mediators; metabolomics thus provides a window into how certain foods might mitigate exercise-induced stress at the molecular level.

Untargeted metabolomics has also been used as a compliance or exposure measure for diets in athletes. For example, metabolomic "fingerprinting" can distinguish athletes consuming a high-protein diet *vs.* high-carbohydrate diet by detecting differences in urea cycle metabolites, amino acid catabolites, and even food-specific compounds (like citrus flavonoid metabolites if one diet is fruit-rich). One study of fencers supplemented with beetroot juice (rich in nitrates) used untargeted metabolomics to monitor the 4 weeks metabolic effect: it found higher circulating nitrite/nitrate and related nitrogenous metabolites in the beetroot group, confirming uptake, as well as enhanced TCA cycle and antioxidant metabolites that correlated with improved time-to-fatigue.[100] In sports anti-doping research, metabolomics is being explored to detect signatures of illicit substances or extreme diets; for instance, certain steroid metabolites or xenobiotics can be picked up in untargeted urine screens and signal potential doping.[101]

Overall, nutritional metabolomics in sports emphasizes individual responses. There is high interest in precision nutrition: tailoring diets to an athlete's metabolic profile. Metabolomics helps define "responders" *vs.* "non-responders." For example, some individuals show a strong metabolite shift after fish oil supplementation, while others show minimal changes – untargeted profiles might help explain such variability by revealing differences in absorption or baseline metabolism.[102,103] In a study of green tea extract, two groups of men had different metabolic responses during exercise despite the same dose: metabolomics indicated that only one subgroup showed significant increases in catechin phase II metabolites and corresponding antioxidative metabolites, suggesting differences in gut microbiome or enzyme activity. These examples illustrate how untargeted metabolomics not only measures the average effect of a nutrition intervention but can uncover inter-individual differences in metabolism that might be crucial for personalized nutrition strategies. As metabolomics datasets grow, machine learning approaches are being tested to predict the optimal diet for a given athlete's metabolic makeup – for instance, using baseline metabolite profiles to predict who will benefit most from carbohydrate loading or from certain supplements.[87]

In summary, applications of untargeted metabolomics in sports nutrition have demonstrated: (1) a robust ability to catalog the biochemical "footprint" of exercise – largely confirming known pathways but also finding new candidate biomarkers of performance and fatigue; (2) the capacity to differentiate metabolic phenotypes of trained *vs.* untrained individuals and monitor training-induced changes; and (3) insights into how specific dietary components influence metabolism before, during, and after exercise. These advances come with the caveat that results must be carefully validated (one study's significant metabolite may not always replicate in another due to methodological differences). Nonetheless, the rich data from untargeted pipelines have undoubtedly expanded our understanding of exercise nutrition biochemistry. The next sections will critically evaluate the strengths and weaknesses of these approaches, and how the field is addressing the controversies that have emerged.

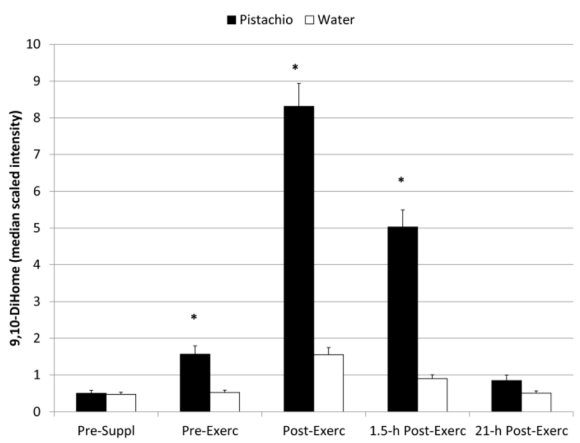

Fig. 4 Plasma 9,10-DiHOME in pistachio and water conditions (interaction effect, $Q < 0.001$) (mean ± SE). Adapted/reproduced from ref. 63 with permission from PLOS,[63] copyright 2014.

General limitations of untargeted metabolomics

Untargeted metabolomics has opened new frontiers in sports nutrition research, but it also faces pointed criticism and

practical hurdles. This section discusses the main pros and cons of untargeted metabolomic pipelines as debated in the field, and pinpoints the focus of ongoing controversies. We also highlight emerging solutions and future directions aimed at maximizing the benefits of untargeted approaches while mitigating their weaknesses (Table 3 provides a summary).

Strengths and advantages

Proponents of untargeted metabolomics argue that its comprehensive scope is uniquely suited for discovery in the complex, integrative context of exercise physiology. One key advantage is the ability to simultaneously capture changes across multiple metabolic pathways without bias. In an exercise or nutrition intervention, dozens of pathways (lipid oxidation, glycolysis, amino acid catabolism, nucleotide turnover, *etc.*) are engaged – an untargeted assay can detect metabolites from all of these, often yielding unexpected findings. For instance, the discovery of an exercise-induced tyrosine metabolism disturbance (mimicking hawkinsinuria) came only because researchers measured broadly and noticed an unusual accumulation of tyrosine byproducts in post-soccer match urine.[104] Targeted assays focused on, say, energy metabolites might have missed this subtle pathway. Thus, untargeted studies can generate new hypotheses about biochemical links between nutrition, metabolism, and performance that were not originally anticipated. Supporters also note that metabolomics captures the phenotype most proximally – metabolites are the end-products of gene and protein activity and reflect real-time physiology.[105] This makes metabolomics a powerful tool for integrating genotype and environment: two athletes might have different gene profiles, but metabolomics shows the net effect of genes, training, and diet combined on their biochemistry. From a practical standpoint, many metabolomics measurements (especially in plasma or urine) are minimally invasive compared to muscle biopsies or organ imaging, enabling more frequent sampling to monitor responses to training or diet changes.[1] The breadth of data also allows for advanced analytics – pattern recognition and machine learning on metabolomic data have shown promise in predicting outcomes like overtraining or risk of injury, earlier than traditional metrics.[106] Moreover, untargeted pipelines encourage data-driven, unbiased exploration, which is valuable in a field like nutrition where we might not even know what the most important molecules are. The discovery of new potential biomarkers (*e.g.* certain acylcarnitine ratios as markers of incomplete fat oxidation, or specific gut microbiome-derived metabolites linked to endurance) stems directly from the open-minded approach of untargeted analysis. Proponents thus view untargeted metabolomics as an indispensable engine for hypothesis generation and a step toward precision sports nutrition, where

Table 3 Summary of advantages and challenges of untargeted metabolomics pipelines in sports nutrition research

Advantages (pros)	Challenges (cons)
Holistic, unbiased coverage: can detect hundreds of metabolites across diverse pathways in one run, enabling discovery of unexpected changes. *E.g.*, revealed novel exercise-induced metabolites (*e.g.* unusual tyrosine catabolites, gut microbiome-derived compounds) beyond predefined targets	Complex data, difficult interpretation: yields large datasets with many unknowns. Metabolite identification is a major bottleneck – a significant fraction of features remain unannotated. Harder to derive clear mechanistic insight when key metabolites are "feature 123@*m*/*z* 256" with unknown identity
Integrative phenotype readout: metabolome reflects combined genetic and environmental influences, providing a direct readout of physiological state. Useful for assessing overall impact of training/diet on the body (closer to performance outcomes than gene or protein changes)	Reproducibility concerns: results can vary with different data processing pipelines or analytical platforms. Lack of standardized protocols means one study's biomarker may not replicate in another study. "Discovery" metabolites often fail validation due to false positives or batch effects
Biomarker discovery for personalized nutrition: identifies candidate biomarkers of performance, fatigue, or nutritional status that targeted approaches might miss. This can enable precision nutrition – tailoring diets based on an individual's metabolomic profile (*e.g.* detecting who is a high lactate producer or who has lipid oxidation deficiencies)	Data overload and multiple comparisons: hundreds of comparisons raise false discovery risk. Without rigorous stats (FDR correction), studies may report too many "significant" changes by chance. Signals can be overwhelmed by noise, requiring cautious interpretation and validation in independent cohorts
Insights into mechanism and recovery: provides time-resolved metabolic profiles to understand exercise recovery and overtraining. Untargeted assays have found metabolic markers of overreaching (*e.g.* elevated acylcarnitine accumulation) and distinct recovery trajectories for different diets. These insights can guide training load adjustments and nutritional recovery strategies	Pre-analytical variability: extremely sensitive to sample handling and timing. Circadian influences, feeding state, storage conditions can all confound results if not strictly controlled. Differences in protocols between studies (*e.g.* fasting *vs.* fed, morning *vs.* afternoon sampling) contribute to inconsistent findings
Hypothesis generation for other omics: untargeted metabolomics findings can drive new hypotheses tested by targeted assays or other omics. For instance, a spike in a specific metabolite might prompt investigation of the enzyme or gene regulating it. This cross-talk enriches the overall understanding of exercise biology	Translational gap: critics note that translating metabolomic changes into practical advice is not straightforward. Many discovered metabolites have unclear relevance to performance or health, making it hard to act on the findings until further research links them to outcomes. Targeted approaches or known markers are seen as more immediately actionable for coaches/athletes
Adaptable to various samples: can be applied to plasma, urine, saliva, muscle extracts, *etc.*, giving flexibility. Non-invasive biofluids (urine, saliva) allow frequent sampling to monitor metabolic fluctuations without burdening athletes. Untargeted assays have even been used in field settings (with portable devices or rapid sampling protocols)	Cost and expertise: untargeted metabolomics experiments are costly (instrument time, data storage) and require specialized bioinformatics expertise. This can limit sample sizes or replication. High cost and complexity mean findings might not be readily validated in large trials, slowing acceptance into practice

an athlete's unique metabolomic profile can inform personalized dietary recommendations.

Challenges and limitations

Critics of untargeted approaches point out several inherent limitations. First and foremost is the difficulty of metabolite identification and biological interpretation. In a typical untargeted experiment, a large proportion of features remain unidentified or putatively annotated (level 3 or 4 identifications).[107] This means researchers might detect a "significant peak" that differentiates supplement A *vs.* B, but if they cannot determine what that metabolite is, the finding has limited practical value. For example, a study might find a feature at *m*/*z* 256 that rises with training – but without identification, one can only speculate about its role. The metabolite ID bottleneck is well-recognized: "Elucidating the chemical structure and origin of unknown significant compounds detected by untargeted approaches remains a bottleneck". This challenge has led some to question whether untargeted data (with many unknowns) can truly advance mechanistic understanding. Efforts like expanding spectral libraries (*e.g.* the Human Metabolome Database now has >200 000 entries) and community initiatives to share MS2 spectra are underway to improve this, but it remains a limiting step.

Another major issue is data overload and processing variability. Untargeted metabolomics generates huge datasets – often thousands of features per sample. How one handles and filters this data can dramatically affect results. Different software pipelines or parameter settings can yield divergent feature lists from the same raw data. One study noted that applying various missing value imputation or normalization methods changed the set of "significant" metabolites considerably.[108] Thus, there is a concern about reproducibility: are the metabolomic biomarkers reported in one paper reproducible by another lab using a different pipeline? The answer has often been "not exactly." This fuels the controversy – skeptics argue that without standardized workflows, untargeted metabolomics findings can be cherry-picked or may contain false positives. The lack of uniform quality control across studies (*e.g.* how to handle batch effects, how to threshold feature detection) makes it hard to compare results or perform meta-analyses, a challenge the field is actively working to address through guidelines and shared protocols.

Pre-analytical and analytical variability further complicate matters. As discussed earlier, small differences in sample collection or handling can produce metabolite differences that swamp the biological effects of interest. For instance, if one study allowed subjects to eat a light breakfast before exercise and another required an overnight fast, their metabolomic "exercise response" profiles could differ simply due to baseline nutrition. The timing of sample collection is another debated point – some argue that many exercise metabolomics studies capture only the immediate post-exercise snapshot, missing the full kinetic picture.[22] If studies are not aligned in timing, their results might conflict. These inconsistencies fuel controversy about which metabolite changes are "real" or universal.

The statistical interpretation of untargeted data is also a point of contention. With hundreds of metabolites tested, the risk of false positives is high if proper corrections (like false discovery rate control) are not applied. Early metabolomics papers sometimes reported dozens of "significant" changes without adjustment for multiple comparisons, which later might not hold up. Now, better practice is in place, but it means often only the top few changes are robust, others hover at the edge of significance. Scholars opposing over-reliance on untargeted data argue that it can generate so many hypotheses that there is a temptation to engage in post-hoc storytelling – finding some pathway among the many that changed and attributing significance to it, while ignoring others (sometimes referred to as the "Texas sharpshooter fallacy"). Prudent studies now use validation: for example, identifying a candidate biomarker in an untargeted run, then using a targeted assay in a follow-up cohort to verify it. Still, few metabolomic biomarkers discovered for exercise (*e.g.* specific acylcarnitine ratios or amino acid derivatives as fatigue markers) have been validated across independent populations, highlighting the challenge of moving from discovery to confirmed application.[109,110]

One more controversy centers on the actionability of untargeted findings in sports nutrition. While untargeted studies generate interesting lists of altered metabolites, coaches and nutritionists ask: how does this translate to improved training or diet plans? For example, discovering that endurance exercise raises metabolite A fourfold is scientifically intriguing, but does it help an athlete? If metabolite A is a marker of fatigue, perhaps yes – one could monitor it to gauge recovery. But if metabolite A's function is unknown, it's unclear what to do with that knowledge. Some critics thus view untargeted metabolomics as too far removed from practical guidance: lots of noise, hard-to-interpret signals, and only incremental gains in understanding. Proponents respond that this is a short-sighted view – today's obscure metabolite could be tomorrow's key to a new nutritional intervention (as lactate once was: once just a "byproduct," now a focus for training strategies). They also note that metabolomics has already suggested practical angles, such as using certain metabolite ratios to detect overtraining before performance drops, or confirming that an antioxidant-rich diet mitigates specific oxidative metabolites after exercise.

Finally, a subtle but important controversy is untargeted *vs.* targeted metabolomics in sports research. Some experts advocate that targeted methods (measuring a set of known, hypothesis-driven metabolites, like specific hormone levels or known ergogenic biomarkers) are more reliable and easier to standardize, and thus more useful for advancing sports nutrition. Untargeted, in their view, should remain in the discovery realm and not be overinterpreted. Others argue that targeted and untargeted are complementary – the untargeted screens find new leads, and targeted follow-ups confirm and quantify them. A convergence is happening where many labs do a broad untargeted scan and then develop targeted assays for the highlights (like a short list of potential biomarkers).

The debates outlined often boil down to a fundamental question: can we trust and effectively use the complex data from

untargeted metabolomics to meaningfully improve sports nutrition and performance? On one side, enthusiasts say yes – these data are revealing hidden aspects of metabolism (like microbiome contributions, novel signaling metabolites) that will form the basis of personalized nutrition and training programs. On the other side, skeptics worry that without greater rigor and standardization, untargeted metabolomics may produce more confusion than clarity, with each study finding its own set of "significant" metabolites and few consistent outcomes across studies. They also emphasize the need to validate findings in terms of actual performance outcomes: it's interesting if metabolite X increases with a supplement, but does that correlate with better endurance or recovery?

Both sides acknowledge certain needs: (1) improved standardization of workflows (from sample handling to data processing) to enhance reproducibility; (2) expanded reference libraries and sharing of metabolomic data to tackle the identification problem – for example, consortia pooling spectral data to identify unknowns that commonly appear in exercise studies; (3) integrative approaches combining metabolomics with other measurements (hormones, proteomics, genomics) to strengthen conclusions – a metabolite change linked with a gene expression change and a performance change is far more convincing than any alone. Indeed, the future may see hybrid pipelines that merge the breadth of untargeted with the focus of targeted: for example, using untargeted data to guide the creation of targeted multi-metabolite panels that can be routinely used for athlete monitoring.

Table 3 succinctly contrasts the main pros and cons with references. It is evident that while untargeted metabolomics has extraordinary exploratory power, it demands careful execution and interpretation. The controversies today are driving improvements: for instance, the argument over reproducibility has led to initiatives for more open data sharing and metabolomics quality control in multi-center trials. As methods mature, we expect the "pros" to strengthen (with more discoveries actually translating to practice) and the "cons" to be mitigated (with better reliability and clarity of data). The next section concludes with forward-looking perspectives on how these pipelines can be refined and integrated into the toolkit of sports nutrition.

Despite the challenges, there is a general optimism that most limitations can be addressed with advancing technology and better study design. For instance, comprehensive reference libraries and improved algorithms (including machine learning for spectral match) are gradually improving the metabolite ID success rate. International efforts like the Metabolomics Quality Assurance and Propagation (MQAP) have been established to improve reproducibility by standardizing protocols and providing quality control samples across labs. In sports metabolomics specifically, consortia are beginning to share data, which will help discern true biological signals from study-specific noise. The controversies have, in a sense, been productive: they highlight where we must focus (*e.g.* ensuring a metabolomics finding has physiological relevance and is not just statistically significant but trivial).

Sports-specific challenges and mitigations

In athletic settings, several constraints are distinct in mechanism or magnitude and warrant separate handling. First, acute and recovery-phase shifts in plasma volume during endurance and heat-stress exercise alter circulating metabolite concentrations independently of production/clearance. We now recommend reporting concentration changes alongside plasma-volume–corrected estimates using the Dill–Costill equation or updated formulations, and providing hemoglobin/hematocrit (or albumin) to document the correction. This is standard in exercise physiology and prevents misattribution of hemoconcentration to true metabolic effects.[111] Second, in-exercise feeding and supplements (*e.g.*, carbohydrate beverages, bananas, nuts, nitrate sources) introduce food-derived xenobiotics and shift lipid-oxidation markers; untargeted profiles will reflect these inputs. We now explicitly require that trials document intake timing/dose and annotate exogenous metabolites in interpretation. For example, metabolomics in 75 km cycling shows clear separation of conditions and attenuated inflammatory/lipid-oxidation perturbations with carbohydrate or banana ingestion *versus* water, underscoring the need to treat nutrition as an experimental factor rather than noise.[112]

Pitch-side saliva and sweat sampling present matrix-specific issues. Salivary flow rate, pH, and oral contaminants influence untargeted signals; post-match studies describe both advantages and limits, emphasizing the need to record flow/pH and to normalize accordingly. For sweat, localized sweat-rate normalization (or area/time-based collection) materially changes biomarker variance structure and should be reported with device/calibration details.[113] Field logistics favor capillary dried-blood-spot (DBS) or volumetric microsampling, which now support metabolomics/lipidomics and lactate-threshold determination in elite cyclists and other athletes. We therefore reference recent DBS/VAMS studies and suggest reporting validation steps (spot volume accuracy, punch location, hematocrit assessment) when using such approaches in athlete monitoring or race settings.[114] Finally, biological covariates unique to or prevalent in sport can dominate untargeted readouts: menstrual-cycle phase and hormonal contraception shift substrate use and lipid metabolism, and low energy availability/REDs alters endocrine and metabolic status. We now advise prespecifying sex-hormone status and cycle phase (or contraceptive class), or adjusting statistically, and screening for REDs where relevant to endurance cohorts.[115]

Conclusions and future perspectives

Untargeted metabolomics has rapidly evolved from a novel experimental technique to a mainstream tool in sports nutrition and exercise science. Over the past decade, it has proven its value by cataloguing the extensive metabolic perturbations caused by exercise and diet, uncovering both expected changes and unanticipated phenomena. This review has highlighted how untargeted pipelines, when carefully applied, can generate

a holistic understanding of the athlete's biochemical status – an understanding that is key to optimizing nutrition for performance and recovery. At the same time, we have examined the legitimate concerns and challenges that accompany the untargeted approach, from data complexity and reproducibility issues to the perennial puzzle of identifying unknown metabolites.

Moving forward, the field is poised to address these challenges head-on. Improved standardization of untargeted metabolomics workflows is underway, with community guidance consolidating best practices for LC-MS based untargeted metabolomics, including reference materials, data-quality review, and identification/annotation criteria; the 2024 mQACC/NIST workshop report outlines a living QA/QC guidance framework intended for broad adoption and continual updates.[116] In parallel, an easily implementable QC protocol for routine monitoring of data quality ("QComics") has been proposed and evaluated in 2024, facilitating transparent and reportable QC in discovery workflows.[39] Collectively, these efforts enhance inter-laboratory comparability and are directly applicable to exercise and sports-nutrition studies. The advancement of analytical technology will also play a role. Beyond gains in sensitivity and mass accuracy, recent multimodal MS environments now enable structural elucidation of lipids and small molecules with electron-activated dissociation and MS imaging, improving sn-position and C=C localization and thereby reducing isomeric ambiguity in untargeted datasets. The 2024 MS-DIAL 5 platform exemplifies this shift, validating structure assignments against standards and NIST SRM 1950 plasma and explicitly supporting IM-MS and DIA-PASEF data types.[117] Such tools, together with curated MS/MS libraries, directly address long-standing identification bottlenecks that have limited translation in sports metabolomics.

One of the most promising trends is the integration of metabolomics with other omes to build a comprehensive picture. This is now exemplified by the NIH MoTrPAC program, which in 2024 published a whole-organism temporal multiomic atlas of endurance-training adaptations, spanning 19 tissues and nine molecular platforms with publicly accessible data resources. These data sets provide a rigorous substrate for validating metabolomic signals and linking them to proteomic, phospho- and epigenomic remodeling relevant to performance and recovery.[118]

In terms of practical applications, untargeted metabolomics is edging closer to real-world use. Recent athlete-specific resources illustrate feasibility at scale and with rigorous QC. For example, a 2025 Scientific Data descriptor provides a plasma metabolomics and lipidomics dataset in trained racewalkers sampled across four recovery time-points, reporting 859 untargeted metabolites plus >800 targeted analytes/lipids, with publicly accessible raw data and methods suitable for secondary analyses and biomarker discovery.[119] Complementary pilot work continues to compare endurance and strength athletes *versus* sedentary controls, helping to operationalize panels relevant to training status and recovery kinetics.[120] Such datasets support the development of practical, validated "wellness" panels derived from untargeted discovery and refined into targeted assays for longitudinal athlete monitoring. Recent trials also help scope personalization claims by reporting effect magnitudes and heterogeneity. In a randomized crossover study (n = 19), 2 weeks pistachio ingestion before a 75 km time trial resulted in a 4.8% slower performance and higher post-exercise 9,10-DiHOME alongside raffinose/sucrose translocation, quantitatively linking specific food components to oxylipin stress biochemistry; these oxylipins would be denoted as level 1–2 identifications depending on standard/MS/MS confirmation.[121] Complementarily, a 4-arm crossover in 20 cyclists reported attenuated 9+13-HODEs and cytokines with carbohydrate intake from bananas or sugar beverages *versus* water, with model validation (R2Y/Q2Y as above), suggesting that recovery-panel thresholds can be anchored to observed fold-change ranges rather than qualitative trends.[122] Together with field-deployable dried-blood-spot workflows, these data support cautious, validated movement toward routine athlete monitoring, while personalization should remain contingent on replicated effect sizes and explicit identification confidence reporting.[114]

Crucially, addressing the controversies has catalyzed a more rigorous approach in recent studies: greater use of validation cohorts and clearer separation of association *versus* causation, supported by field-level guidance synthesized for exercise researchers in 2024,[123] which we now cite alongside the above QA/QC and software advances. For instance, metabolomics alone might show that a certain dipeptide increases after training; collaboration with biochemists and physiologists could then identify it as a marker of muscle protein turnover, giving it meaning. We see a future where metabolomic data will not exist in silos but will be part of a broader matrix including endocrinology, traditional biomarkers, and performance metrics. This will diminish the chance of overhyping any single metabolite and instead allow metabolomics to enrich our overall understanding of the athlete's adaptive landscape.

In conclusion, evaluating untargeted metabolomics pipelines – as we have done in this review – demonstrates that despite some drawbacks, these approaches are immensely valuable for sports nutrition research. The pros clearly offer transformative opportunities for personalized and optimized nutrition strategies in sport. The cons are being actively mitigated through technical, computational, and collaborative advancements. The focus of controversy has sharpened the science, pushing the community toward higher standards of evidence. Untargeted metabolomics has, in a short time, advanced from simply cataloguing "molecules that change with exercise" to generating targeted hypotheses about improving athlete health and performance – for example, suggesting nutritional interventions to influence specific metabolic pathways. Emerging evidence already quantifies the monitoring potential and begins to define where personalization may be justified. In elite cyclists, dried-blood-spot field sampling captured exertion signatures with two- to three-fold increases in lactate and succinate during graded tests, and larger fatty acid/acylcarnitine responses during long aerobic sessions; comparable signatures appeared during race sprints and climbs, supporting translational monitoring feasibility (n = 28 lab; n =

5 in-race).[114] In a randomized crossover trial of 20 trained cyclists, banana or sugar beverages (*vs.* water) reduced post-exercise inflammatory and lipid-oxidation perturbations; multivariate modeling showed clear separation of treatment arms (OPLS-DA R2Y = 0.848; Q2Y = 0.409; permutation validation passed), with 109 metabolites increasing >2-fold and 71 decreasing <0.5-fold in the water condition, indicating quantifiable, diet-responsive recovery trajectories suitable for compliance and load tracking.[122] Where named metabolites are confirmed against standards (*e.g.*, lactate, succinate) we indicate identification at MSI/Schymanski level 1; library-matched acylcarnitines and other features are flagged as level 2, to make identification confidence explicit for any proposed monitoring panel.[124] In the ongoing quest to understand how the human body responds to exercise and diet, untargeted metabolomics has proven itself as a powerful lens, and with continual refinements, it is poised to bring the blurry edges of our knowledge into ever sharper focus, benefiting both scientific insight and athletic practice.

Author contributions

Yongfu Liu: conceptualization, supervision, writing – review & editing, funding acquisition. Yuting Hu: literature search, data curation, writing – original draft, visualization. Wenjun Yu: methodology, validation, writing – review & editing, visualization.

Conflicts of interest

There are no conflicts to declare.

Data availability

No primary research results, software or code have been included and no new data were generated or analysed as part of this review.

Acknowledgements

This research was funded by the Guided Scientific Research Project of Shiyan Science and Technology Bureau (24Y055).

References

1 S. Qi, X. Li, J. Yu and L. Yin, *Front. Physiol.*, 2024, **14**, 1332104.
2 V. Guntur, T. Nemkov, E. Boer, M. Mohning, D. Baraghoshi, F. Cendali, I. San-Millán, I. Petrache and A. D'Alessandro, *Metabolites*, 2022, **12**, 1026.
3 C. Sands, M. Gómez-Romero, G. Correia, E. Chekmeneva, S. Camuzeaux, C. Izzi-Engbeaya, W. Dhillo, Z. Takáts and M. Lewis, *Anal. Chem.*, 2021, **93**, 1924–1933.
4 M. Spanakis, P. Fragkiadaki, E. Renieri, E. Vakonaki, I. Fragkiadoulaki, A. Alegakis, M. Kiriakakis, N. Panagiotou, E. Ntoumou, I. Gratsias, E. Zoubaneas, G. Morozova, M. Ovchinnikova, C. Tsitsimpikou, K. Tsarouhas, N. Drakoulis, A. Skalny and A. Tsatsakis, *Frontiers in Sports and Active Living*, 2024, **5**, 1327792.
5 R. Muniz-Santos, A. Magno-França, I. Jurišica and L. Cameron, *OMICS: J. Integr. Biol.*, 2023, **27**, 499–518.
6 T. C. L. França, R. Muniz-Santos, L. C. Caetano, G. H. M. F. Souza, H. F. Goulart, M. Assis, A. Bottino, A. Bassini, A. E. G. Santana, E. S. Prado and L. C. Cameron, *Front. Nutr.*, 2023, **10**, 1169188.
7 X. Yu, J. Xiao, S. Chen, Y. Yu, J. Ma, Y. Lin, R. Li, J. Lin, Z. Fu, Q. Zhou, Q. Chao, L. Chen, Z. Yang and R. Liu, *Nat. Commun.*, 2022, **11**, 5586.
8 M. Yu, G. Dolios and L. Petrick, *J. Cheminf.*, 2022, **14**, 6.
9 Z. Pang, L. Xu, C. Viau, Y. Lü, R. Salavati, N. Basu and J. Xia, *Nat. Commun.*, 2024, **15**, 3675.
10 M. Jeppesen and R. Powers, *Magn. Reson. Chem.*, 2023, **61**, 628–653.
11 B. Hoegen, J. Hampstead, U. Engelke, P. Kulkarni, R. Wevers, H. Brunner, K. Coene and C. Gilissen, *J. Inherited Metab. Dis.*, 2022, **45**, 682–695.
12 A. Cabrita, J. Guilherme-Fernandes, I. Valente, A. Almeida, S. Lima, A. Fonseca and M. Maia, *Animals*, 2022, **12**, 2643.
13 Z. Zhou, M. Luo, X. Chen, Y. Yin, X. Xiong, R. Wang and Z. Zhu, *Nat. Commun.*, 2020, **11**, 4334.
14 G. J. Gouveia, A. O. Shaver, B. M. Garcia, A. M. Morse, E. C. Andersen, A. S. Edison and L. M. McIntyre, *Anal. Chem.*, 2021, **93**(26), 9193–9199.
15 L. Litichevskiy and C. Thaiss, *Annu. Rev. Nutr.*, 2022, **42**, 145–164.
16 H. Nguyen, F. Hu, R. Reisdorph, K. Doenges, E. Willis, M. Breit, J. Broussard, V. Catenacci, S. Creasy and S. Borengasser, *Diabetes*, 2025, **74**, 1908–LB.
17 P. Dimofski, D. Meyre, N. Dreumont and B. Leininger-Muller, *Nutrients*, 2021, **13**, 2818.
18 F. Alsenani, K. Tupally, E. Chua, E. Eltanahy, H. Alsufyani, H. Parekh and P. Schenk, *Saudi Pharm. J.*, 2020, **28**, 1834–1841.
19 Y. Zhang, Y. Shen, W. Cheng, X. Wang, Y. Xue, X. Chen and B. Han, *Front. Microbiol.*, 2021, **12**, 714726.
20 A. Maroli and R. Powers, *NMR Biomed.*, 2023, **36**, e4594.
21 C. Chen, D. Lee, J. Yu, Y. Lin and T. Lin, *Mass Spectrom. Rev.*, 2022, **42**, 2349–2378.
22 C. A. Sakaguchi, D. C. Nieman, E. F. Signini, R. M. Abreu and A. M. Catai, *Metabolites*, 2019, **9**, 164.
23 A. Sens, S. Rischke, L. Hahnefeld, E. Dorochow, S. Schäfer, D. Thomas, M. Köhm, G. Geißlinger, F. Behrens and R. Gurke, *J. Mass Spectrom. Adv. Clin. Lab*, 2023, **28**, 35–46.
24 X. Huang, Y. Wang, W. Wang, X. Zhang, L. Jiang, J. Liu, S. Liu, K. Li, C. Xie and Q. Wang, *Anal. Sci.*, 2022, **39**, 221–227.
25 C. Stoffels, T. Angerer, H. Robert, N. Poupin, L. Lakhal, G. Frache, M. Mercier-Bonin and J. Audinot, *Anal. Chem.*, 2023, **95**, 6568–6576.
26 K. Contrepois, S. Wu, K. J. Moneghetti, D. Hornburg, S. Ahadi, M.-S. Tsai, A. A. Metwally, E. Wei, B. Lee-McMullen, J. V. Quijada, S. Chen, J. W. Christle, M. Ellenberger, B. Balliu, S. Taylor, M. G. Durrant, D. A. Knowles, H. Choudhry, M. Ashland, A. Bahmani,

B. Enslen, M. Amsallem, Y. Kobayashi, M. Avina, D. Perelman, S. M. S.-F. Rose, W. Zhou, E. A. Ashley, S. B. Montgomery, H. Chaib, F. Haddad and M. P. Snyder, *Cell*, 2020, **181**, 1112–1130.

27 S. W. Harshman, R. L. Pitsch, Z. K. Smith, M. L. O'Connor, B. A. Geier, A. V. Qualley, N. M. Schaeublin, M. V. Fischer, J. J. Eckerle, A. J. Strang and J. A. Martin, *PLoS One*, 2018, **13**, e0203133.

28 J. King, P. Mochalski, A. Kupferthaler, K. Unterkofler, H. Koc, W. Filipiak, S. Teschl, H. Hinterhuber and A. Amann, *Physiological Measurement*, 2010, **31**, 1169.

29 N. Jonge, K. Mildau, D. Meijer, J. Louwen, C. Bueschl, F. Huber and J. Hooft, *Metabolomics*, 2022, **18**, 103.

30 J. Guo and T. Huan, *Anal. Chem.*, 2023, **95**, 5894–5902.

31 C. Yang, A. Zhang, Z. Gao and G. Su, *Chin. J. Chromatogr.*, 2025, **43**, 585–593.

32 Z. Guo, Y. Fan, C. Yu, H. Lü and Z. Zhang, *Anal. Chem.*, 2024, **96**, 5878–5886.

33 G. Davison, M. Vinaixa, R. McGovern, A. Beltran, A. Novials, X. Correig and C. McClean, *Front. Physiol.*, 2018, **9**, 1682.

34 M. R. Belhaj, N. G. Lawler and N. J. Hoffman, *Metabolites*, 2021, **11**, 151.

35 E. E. Kontou, A. Walter, O. Alka, J. Pfeuffer, T. Sachsenberg, O. S. Mohite, M. Nuhamunada, O. Kohlbacher and T. Weber, *J. Cheminf.*, 2023, **15**, 52.

36 R. Tautenhahn, C. Böttcher and S. Neumann, *BMC Bioinf.*, 2008, **9**, 504.

37 M. Aigensberger, C. Bueschl, E. Castillo-Lopez, S. Ricci, R. Rivera-Chacon, Q. Zebeli, F. Berthiller and H. E. Schwartz-Zimmermann, *Anal. Chim. Acta*, 2025, **1336**, 343491.

38 H. Habra, M. Kachman, K. Bullock, C. Clish, C. R. Evans and A. Karnovsky, *Anal. Chem.*, 2021, **93**, 5028–5036.

39 Á. González-Domínguez, N. Estanyol-Torres, C. Brunius, R. Landberg and R. González-Domínguez, *Anal. Chem.*, 2024, **96**, 1064–1072.

40 C. D. Broeckling, R. D. Beger, L. L. Cheng, R. Cumeras, D. J. Cuthbertson, S. Dasari, W. C. Davis, W. B. Dunn, A. M. Evans, A. Fernández-Ochoa, H. Gika, R. Goodacre, K. D. Goodman, G. J. Gouveia, P.-C. Hsu, J. A. Kirwan, D. Kodra, J. Kuligowski, R. S.-L. Lan, M. E. Monge, L. W. Moussa, S. G. Nair, N. Reisdorph, S. D. Sherrod, C. Ulmer Holland, D. Vuckovic, L.-R. Yu, B. Zhang, G. Theodoridis and J. D. Mosley, *Anal. Chem.*, 2023, **95**, 18645–18654.

41 J. Ivanisevic and E. J. Want, *Metabolites*, 2019, **9**, 308.

42 B. Li, J. Tang, Q. Yang, S. Li, X. Cui, Y. Li, Y. Chen, W. Xue, X. Li and F. Zhu, *Nucleic Acids Res.*, 2017, **45**, W162–W170.

43 P. Stincone, A. K. Pakkir Shah, R. Schmid, L. G. Graves, S. P. Lambidis, R. R. Torres, S.-N. Xia, V. Minda, A. T. Aron, M. Wang, C. C. Hughes and D. Petras, *Anal. Chem.*, 2023, **95**, 12673–12682.

44 L. W. Sumner, A. Amberg, D. Barrett, M. H. Beale, R. Beger, C. A. Daykin, T. W.-M. Fan, O. Fiehn, R. Goodacre, J. L. Griffin, T. Hankemeier, N. Hardy, J. Harnly, R. Higashi, J. Kopka, A. N. Lane, J. C. Lindon, P. Marriott, A. W. Nicholls, M. D. Reily, J. J. Thaden and M. R. Viant, *Metabolomics*, 2007, **3**, 211–221.

45 R. Schmid, S. Heuckeroth, A. Korf, A. Smirnov, O. Myers, T. S. Dyrlund, R. Bushuiev, K. J. Murray, N. Hoffmann, M. Lu, A. Sarvepalli, Z. Zhang, M. Fleischauer, K. Dührkop, M. Wesner, S. J. Hoogstra, E. Rudt, O. Mokshyna, C. Brungs, K. Ponomarov, L. Mutabdžija, T. Damiani, C. J. Pudney, M. Earll, P. O. Helmer, T. R. Fallon, T. Schulze, A. Rivas-Ubach, A. Bilbao, H. Richter, L.-F. Nothias, M. Wang, M. Orešič, J.-K. Weng, S. Böcker, A. Jeibmann, H. Hayen, U. Karst, P. C. Dorrestein, D. Petras, X. Du and T. Pluskal, *Nat. Biotechnol.*, 2023, **41**, 447–449.

46 H. Takeda, Y. Matsuzawa, M. Takeuchi, M. Takahashi, K. Nishida, T. Harayama, Y. Todoroki, K. Shimizu, N. Sakamoto, T. Oka, M. Maekawa, M. H. Chung, Y. Kurizaki, S. Kiuchi, K. Tokiyoshi, B. Buyantogtokh, M. Kurata, A. Kvasnička, U. Takeda, H. Uchino, M. Hasegawa, J. Miyamoto, K. Tanabe, S. Takeda, T. Mori, R. Kumakubo, T. Tanaka, T. Yoshino, M. Okamoto, H. Takahashi, M. Arita and H. Tsugawa, *Nat. Commun.*, 2024, **15**, 9903.

47 J. Pfeuffer, C. Bielow, S. Wein, K. Jeong, E. Netz, A. Walter, O. Alka, L. Nilse, P. D. Colaianni, D. McCloskey, J. Kim, G. Rosenberger, L. Bichmann, M. Walzer, J. Veit, B. Boudaud, M. Bernt, N. Patikas, M. Pilz, M. P. Startek, S. Kutuzova, L. Heumos, J. Charkow, J. C. Sing, A. Feroz, A. Siraj, H. Weisser, T. M. H. Dijkstra, Y. Perez-Riverol, H. Röst, O. Kohlbacher and T. Sachsenberg, *Nat. Methods*, 2024, **21**, 365–367.

48 Z. Pang, G. Zhou, J. Ewald, L. Chang, O. Hacariz, N. Basu and J. Xia, *Nat. Protoc.*, 2022, **17**, 1735–1761.

49 H. Luan, F. Ji, Y. Chen and Z. Cai, *Anal. Chim. Acta*, 2018, **1036**, 66–72.

50 M. Hoffmann, L. Nothias, M. Ludwig, M. Fleischauer, E. Gentry, M. Witting, P. Dorrestein, K. Dührkop and S. Böcker, *Nat. Biotechnol.*, 2021, **40**, 411–421.

51 L.-F. Nothias, D. Petras, R. Schmid, K. Dührkop, J. Rainer, A. Sarvepalli, I. Protsyuk, M. Ernst, H. Tsugawa, M. Fleischauer, F. Aicheler, A. A. Aksenov, O. Alka, P.-M. Allard, A. Barsch, X. Cachet, A. M. Caraballo-Rodriguez, R. R. Da Silva, T. Dang, N. Garg, J. M. Gauglitz, A. Gurevich, G. Isaac, A. K. Jarmusch, Z. Kameník, K. B. Kang, N. Kessler, I. Koester, A. Korf, A. Le Gouellec, M. Ludwig, H. C. Martin, L.-I. McCall, J. McSayles, S. W. Meyer, H. Mohimani, M. Morsy, O. Moyne, S. Neumann, H. Neuweger, N. H. Nguyen, M. Nothias-Esposito, J. Paolini, V. V. Phelan, T. Pluskal, R. A. Quinn, S. Rogers, B. Shrestha, A. Tripathi, J. J. J. van der Hooft, F. Vargas, K. C. Weldon, M. Witting, H. Yang, Z. Zhang, F. Zubeil, O. Kohlbacher, S. Böcker, T. Alexandrov, N. Bandeira, M. Wang and P. C. Dorrestein, *Nat. Methods*, 2020, **17**, 905–908.

52 G. Lefort, L. Liaubet, C. Canlet, P. Tardivel, M.-C. Père, H. Quesnel, A. Paris, N. Iannuccelli, N. Vialaneix and R. Servien, *Bioinformatics*, 2019, **35**, 4356–4363.

53 J. Hao, W. Astle, M. De Iorio and T. M. D. Ebbels, *Bioinformatics*, 2012, **28**, 2088–2090.
54 M. Salem, A. Zayed, S. Alseekh, A. Fernie and P. Giavalisco, *Phytochemistry*, 2021, **190**, 112843.
55 J. A. Westerhuis, H. C. J. Hoefsloot, S. Smit, D. J. Vis, A. K. Smilde, E. J. J. van Velzen, J. P. M. van Duijnhoven and F. A. van Dorsten, *Metabolomics*, 2008, **4**, 81–89.
56 B. Worley and R. Powers, *Curr. Metabolomics*, **1**, 92–107.
57 Multilevel – mixOmics, **https://mixomics.org/methods/multilevel/**, accessed October 8, 2025.
58 T. Newans, P. Bellinger, C. Drovandi, S. Buxton and C. Minahan, *International Journal of Sports Physiology and Performance*, 2022, **17**(8), 1289–1295.
59 S. Das, Y. Kumar, S. Sharma, R. Ray, S. Arava, S. Seth, A. Agarwal and G. Sharma, *Sci. Rep.*, 2022, **12**, 5238.
60 S. Zuffa, R. Schmid, A. Bauermeister, P. Gomes, A. Caraballo-Rodríguez, Y. Abiead, A. Aron, E. Gentry, J. Zemlin, M. Meehan, N. Avalon, R. Cichewicz, E. Buzun, M. Terrazas, C. Hsu, R. Oles, A. Ayala, J. Zhao, H. Chu, M. Kuijpers, S. Jackrel, F. Tugizimana, L. Nephali, I. Dubery, N. Madala, E. Moreira, L. Costa-Lotufo, N. Lopes, P. Rezende-Teixeira, P. Jimenez, B. Rimal, A. Patterson, M. Traxler, R. Pessotti, D. Alvarado-Villalobos, G. Tamayo–Castillo, P. Chaverrí, E. Escudero-Leyva, L. Quirós-Guerrero, A. Bory, J. Joubert, A. Rutz, J. Wolfender, P. Allard, A. Sichert, S. Pontrelli, B. Pullman, N. Bandeira, W. Gerwick, K. Gindro, J. Massana-Codina, B. Wagner, K. Forchhammer, D. Petras, N. Aiosa, N. Garg, M. Liebeke, P. Bourceau, K. Kang, H. Gadhavi, L. Carvalho, M. Santos, A. Pérez-Lorente, C. Molina-Santiago, D. Romero, R. Franke, M. Brönstrup, A. León, P. Pope, S. Rosa, G. Barbera, H. Roager, M. Laursen, F. Hammerle, B. Siewert, U. Peintner, C. Licona-Cassani, L. Rodríguez-Orduña, E. Rampler, F. Hildebrand, G. Koellensperger, H. Schoeny, K. Hohenwallner, L. Panzenboeck, R. Gregor, E. O'Neill, E. Roxborough, J. Odoi, N. Bale, S. Ding, J. Damsté, X. Guan, J. Cui, K. Ju, D. Silva, F. Silva, G. Silva, H. Koolen, C. Grundmann and J. Clement, *Nat. Microbiol.*, 2024, **9**, 336–345.
61 F. Wang, D. Allen, S. Tian, E. Oler, V. Gautam, R. Greiner, T. Metz and D. Wishart, *Nucleic Acids Res.*, 2022, **50**, W165–W174.
62 Y. Li, W. Zhu, Q. Xiang, J. Kim, C. Dufresne, Y. Liu, T. Li and S. Chen, *Int. J. Mol. Sci.*, 2023, **24**, 2249.
63 D. C. Nieman, J. Scherr, B. Luo, M. P. Meaney, D. Dréau, W. Sha, D. A. Dew, D. A. Henson and K. L. Pappan, *PLoS One*, 2014, **9**, e113725.
64 D. Wishart, A. Guo, E. Oler, F. Wang, A. Anjum, H. Peters, R. Dizon, Z. Sayeeda, S. Tian, B. Lee, M. Berjanskii, R. Mah, M. Yamamoto, J. Jovel, C. Torres-Calzada, M. Hiebert-Giesbrecht, V. Lui, D. Varshavi, D. Varshavi, D. Allen, D. Arndt, N. Khetarpal, A. Sivakumaran, K. Harford, S. Sanford, K. Yee, X. Cao, Z. Budinski, J. Liigand, L. Zhang, J. Zheng, R. Mandal, N. Karu, M. Dambrova, H. Schiöth, R. Greiner and V. Gautam, *Nucleic Acids Res.*, 2021, **50**, D622–D631.
65 T. Hovinen, L. Korkalo, R. Freese, E. Skaffari, P. Isohanni, M. Niemi, J. Nevalainen, H. Gylling, N. Zamboni, M. Erkkola and A. Suomalainen, *EMBO Mol. Med.*, 2021, **13**, e13492.
66 Z. Yang, D. Wang, Y. Li, X. Zhou, T. Liu, C. Shi, R. Li, Y. Zhang, J. Zhang, J. Yan, X. Zhu, Y. Li, M. Gong, C. Wang, C. Yuan, Y. Cui and X. Wu, *Biomed. Pharmacother.*, 2021, **146**, 112495.
67 V. Aya, D. Pardo-Rodriguez, L. C. Vega, M. P. Cala and J. D. Ramírez, *Sci. Rep.*, 2025, **15**, 15356.
68 L. Alvares, R. Ferreira, F. Nakamoto, M. Quaresma, L. Santos, L. Degani-Costa, L. Navarro, G. Navarro, A. Oliveira-Junior, R. Barbosa and F. Lima, *BMJ Open Sport & Exercise Medicine*, 2023, **9**, e001641.
69 G. Li, P. Huang, S. Cui, Y. Tan, Y. He, X. Shen, Q. Jiang, P. Huang, G. He, B. Li, Y. Li, J. Xu, Z. Wang and S. Chen, *Transl. Neurodegener.*, 2022, **11**, 6.
70 T. Bekfani, M. Bekhite, S. Neugebauer, S. Derlien, A. Hamadanchi, J. Nisser, M. Hilse, D. Haase, T. Kretzschmar, M. Wu, M. Lichtenauer, M. Kiehntopf, S. Haehling, P. Schlattmann, G. Lehmann, M. Franz, S. Möbius-Winkler and C. Schulze, *Cells*, 2022, **11**, 1674.
71 S. Sutehall and Y. Pitsiladis, *Scandinavian Journal of Medicine and Science in Sports*, 2025, **35**, e70044.
72 M. Nayor, R. Shah, P. Miller, J. Blodgett, M. Tanguay, A. Pico, V. Murthy, R. Malhotra, N. Houstis, A. Deik, K. Pierce, K. Bullock, L. Dailey, R. Velagaleti, S. Moore, J. Ho, A. Baggish, C. Clish, M. Larson, R. Vasan and G. Lewis, *Circulation*, 2020, **142**, 1905–1924.
73 Y. Zou, Y. Wang, H. Ma, Z. He, Y. Tang, L. Guo, Y. Liu, M. Ding, S. Qian and Q. Tang, *J. Lipid Res.*, 2020, **61**, 1589–1604.
74 J. García-Gimenez, I. Cánovas-Cervera and F. Pallardó, *Free Radicals Biol. Med.*, 2024, **213**, 123–137.
75 T. Gibbons, J. Cotter, P. Ainslie, W. Abraham, B. Mockett, H. Campbell, E. Jones, E. Jenkins and K. Thomas, *J. Physiol.*, 2023, **601**, 2121–2137.
76 R. Robergs, K. Leo, L. Mulligan, G. Nygaard, J. Davies, T. Clarke and J. Holland, *Qeios*, 2024, **6**, 1.
77 H. Taha, *J. Physiol.*, 2024, **602**, 6657–6658.
78 A. B. Hodgson, R. K. Randell, N. Boon, U. Garczarek, D. J. Mela, A. E. Jeukendrup and D. M. Jacobs, *J. Nutr. Biochem.*, 2013, **24**, 325–334.
79 D. C. Nieman, N. D. Gillitt, A. M. Knab, R. A. Shanely, K. L. Pappan, F. Jin and M. A. Lila, *PLoS One*, 2013, **8**, e72215.
80 D. C. Nieman, N. D. Gillitt, W. Sha, M. P. Meaney, C. John, K. L. Pappan and J. M. Kinchen, *J. Proteome Res.*, 2015, **14**, 5367–5377.
81 J. P. Karl, L. M. Margolis, N. E. Murphy, C. T. Carrigan, J. W. Castellani, E. H. Madslien, H.-K. Teien, S. Martini, S. J. Montain and S. M. Pasiakos, *Physiol. Rep.*, 2017, **5**, e13407.
82 F. A. Manaf, N. G. Lawler, J. J. Peiffer, G. L. Maker, M. C. Boyce, T. J. Fairchild and D. Broadhurst, *J. Appl. Physiol.*, 2018, **125**, 1193–1203.

83 U. M. Kujala, V.-P. Mäkinen, I. Heinonen, P. Soininen, A. J. Kangas, T. H. Leskinen, P. Rahkila, P. Würtz, V. Kovanen, S. Cheng, S. Sipilä, M. Hirvensalo, R. Telama, T. Tammelin, M. J. Savolainen, A. Pouta, P. F. O'Reilly, P. Mäntyselkä, J. Viikari, M. Kähönen, T. Lehtimäki, P. Elliott, M. J. Vanhala, O. T. Raitakari, M.-R. Järvelin, J. Kaprio, H. Kainulainen and M. Ala-Korpela, *Circulation*, 2013, **127**, 340–348.
84 G. Blackburn, J. Hay, C. Skagen, E. Paul, F. Achcar, J. Wilson, C. Best, E. Manson, K. Burgess, M. Barrett and J. Gill, *Metabolites*, 2020, **10**, 399.
85 D. Fernández-Lázaro, S. Arribalzaga, E. Gutiérrez-Abejón, M. Azarbayjani, J. Mielgo-Ayuso and E. Roche, *Nutrients*, 2024, **16**, 2044.
86 C. Sun, A. Wang, Y. Zhou, P. Chen, X. Wang, J. Huang, J. Gao, X. Wang, L. Shu, J. Lu, W. Dai, Z. Bu, J. Ji and J. He, *Nat. Commun.*, 2023, **14**, 2692.
87 D. M. Jacobs, A. B. Hodgson, R. K. Randell, K. Mahabir-Jagessar-T, U. Garczarek, A. E. Jeukendrup, D. J. Mela and S. Lotito, *J. Agric. Food Chem.*, 2014, **62**, 9936–9943.
88 F. Al-Khelaifi, I. Diboun, F. Donati, F. Botrè, M. Alsayrafi, C. Georgakopoulos, K. Suhre, N. A. Yousri and M. A. Elrayess, *Sports Medicine - Open*, 2018, **4**, 2.
89 C. M. Neal, A. M. Hunter, L. Brennan, A. O'Sullivan, D. L. Hamilton, G. DeVito and S. D. R. Galloway, *J. Appl. Physiol.*, 2013, **114**, 461–471.
90 M. Sellami, M. Elrayess, L. Puce and N. Bragazzi, *Front. Mol. Biosci.*, 2022, **8**, 815410.
91 M. Palmnäs, C. Brunius, L. Shi, A. Rostgaard-Hansen, N. E. Torres, R. González-Domínguez, R. Zamora-Ros, Y. L. Ye, J. Halkjær, A. Tjønneland, G. Riccardi, R. Giacco, G. Costabile, C. Vetrani, J. Nielsen, C. Andres-Lacueva and R. Landberg, *Adv. Nutr.*, 2020, **11**, 524–532.
92 A. Riedl, C. Gieger, H. Hauner, H. Daniel and J. Linseisen, *Br. J. Nutr.*, 2017, **117**, 1631–1644.
93 F. Al-Khelaifi, I. Diboun, F. Donati, F. Botrè, M. Alsayrafi, C. Georgakopoulos, K. Suhre, N. A. Yousri and M. A. Elrayess, *Sports Medicine - Open*, 2018, **4**, 2.
94 M. Parstorfer, G. Poschet, D. Kronsteiner, K. Brüning and B. Friedmann-Bette, *Metabolites*, 2023, **13**, 833.
95 Ø. Karlsson, A. D. Govus, K. McGawley and H. G. Hanstock, *Sports Medicine - Open*, 2024, **10**, 99.
96 R. Shi, J. Zhang, B. Fang, X. Tian, Y. Feng, Z. Cheng, Z. Fu, J. Zhang and J. Wu, *Nutr. Metab.*, 2020, **17**, 19.
97 J. F. Schader, M. Haid, A. Cecil, J. Schoenfeld, M. Halle, A. Pfeufer, C. Prehn, J. Adamski, D. C. Nieman and J. Scherr, *Metabolites*, 2020, **10**, 87.
98 R. S. Kelly, M. P. Kelly and P. Kelly, *Biochim. Biophys. Acta, Mol. Basis Dis.*, 2020, **1866**, 165936.
99 M. Mietus-Snyder, N. Narayanan, R. Krauss, K. Laine-Graves, J. McCann, M. Shigenaga, T. McHugh, B. Ames and J. Suh, *PLoS One*, 2020, **15**, e0240437.
100 C. Zhao, Y. Gong, L. Zheng and M. Zhao, *Food Res. Int.*, 2024, **184**, 114261.
101 B. Keen, A. Cawley, B. Reedy and S. Fu, *Drug Test. Anal.*, 2022, **14**, 794–807.
102 C. B. O'Donovan, M. C. Walsh, M. J. Gibney, E. R. Gibney and L. Brennan, *Proc. Nutr. Soc.*, 2016, **75**, 106–114.
103 E. Holmes, I. D. Wilson and J. K. Nicholson, *Cell*, 2008, **134**, 714–717.
104 D. J. Wilkinson, N. J. Smeeton and P. W. Watt, *Prog. Neurobiol.*, 2010, **91**, 200–219.
105 A. Tebani and S. Bekri, *Front. Nutr.*, 2019, **6**, 41.
106 L. M. Sparks, *Diabetologia*, 2017, **60**, 2329–2336.
107 D. González-Peña and L. Brennan, *Annu. Rev. Food Sci. Technol.*, 2019, **10**, 479–519.
108 P. S. Larmo, A. J. Kangas, P. Soininen, H.-M. Lehtonen, J.-P. Suomela, B. Yang, J. Viikari, M. Ala-Korpela and H. P. Kallio, *Am. J. Clin. Nutr.*, 2013, **98**, 941–951.
109 L. M. Heaney, K. Deighton and T. Suzuki, *Journal of Sports Sciences*, 2019, **37**, 959–967.
110 G. D. Lewis, L. Farrell, M. J. Wood, M. Martinovic, Z. Arany, G. C. Rowe, A. Souza, S. Cheng, E. L. McCabe, E. Yang, X. Shi, R. Deo, F. P. Roth, A. Asnani, E. P. Rhee, D. M. Systrom, M. J. Semigran, R. S. Vasan, S. A. Carr, T. J. Wang, M. S. Sabatine, C. B. Clish and R. E. Gerszten, *Sci. Transl. Med.*, 2010, **2**, 33ra37.
111 D. B. Dill and D. L. Costill, *J. Appl. Physiol.*, 1974, **37**, 247–248.
112 D. C. Nieman, N. D. Gillitt, D. A. Henson, W. Sha, R. A. Shanely, A. M. Knab, L. Cialdella-Kam and F. Jin, *PLoS One*, 2012, **7**, e37479.
113 E. Pitti, G. Petrella, S. Di Marino, V. Summa, M. Perrone, S. D'Ottavio, A. Bernardini and D. O. Cicero, *Metabolites*, 2019, **9**, 141.
114 T. Nemkov, F. Cendali, D. Stefanoni, J. L. Martinez, K. C. Hansen, I. San-Millán and A. D'Alessandro, *Sports Med.*, 2023, **53**, 1651–1665.
115 M. A. Carmichael, R. L. Thomson, L. J. Moran and T. P. Wycherley, *Int. J. Environ. Res. Public Health*, 2021, **18**, 1667.
116 T. Schock, J. Mosley, W. Dunn, J. Kuligowski, M. Lewis, M. Monge, C. U. Holland, D. Vuckovic and K. Zanetti, *Metabolomics*, 2024, **20**, 73.
117 H. Takeda, Y. Matsuzawa, M. Takeuchi, M. Takahashi, K. Nishida, T. Harayama, Y. Todoroki, K. Shimizu, N. Sakamoto, T. Oka, M. Maekawa, M. H. Chung, Y. Kurizaki, S. Kiuchi, K. Tokiyoshi, B. Buyantogtokh, M. Kurata, A. Kvasnička, U. Takeda, H. Uchino, M. Hasegawa, J. Miyamoto, K. Tanabe, S. Takeda, T. Mori, R. Kumakubo, T. Tanaka, T. Yoshino, M. Okamoto, H. Takahashi, M. Arita and H. Tsugawa, *Nat. Commun.*, 2024, **15**, 9903.
118 D. Amar, N. R. Gay, P. M. Jean-Beltran, D. Bae, S. Dasari, C. Dennis, C. R. Evans, D. A. Gaul, O. Ilkayeva, A. A. Ivanova, M. T. Kachman, H. Keshishian, I. R. Lanza, A. C. Lira, M. J. Muehlbauer, V. D. Nair, P. D. Piehowski, J. L. Rooney, K. S. Smith, C. L. Stowe, B. Zhao, N. M. Clark, D. Jimenez-Morales, M. E. Lindholm, G. M. Many, J. A. Sanford, G. R. Smith, N. G. Vetr, T. Zhang, J. J. Almagro Armenteros, J. Avila-Pacheco, N. Bararpour, Y. Ge, Z. Hou, S. Marwaha, D. M. Presby, A. Natarajan Raja, E. M. Savage, A. Steep, Y. Sun, S. Wu,

J. Zhen, S. C. Bodine, K. A. Esser, L. J. Goodyear, S. Schenk, S. B. Montgomery, F. M. Fernández, S. C. Sealfon, M. P. Snyder, J. N. Adkins, E. Ashley, C. F. Burant, S. A. Carr, C. B. Clish, G. Cutter, R. E. Gerszten, W. E. Kraus, J. Z. Li, M. E. Miller, K. S. Nair, C. Newgard, E. A. Ortlund, W.-J. Qian, R. Tracy, M. J. Walsh, M. T. Wheeler, K. P. Dalton, T. Hastie, S. G. Hershman, M. Samdarshi, C. Teng, R. Tibshirani, E. Cornell, N. Gagne, S. May, B. Bouverat, C. Leeuwenburgh, C. Lu, M. Pahor, F.-C. Hsu, S. Rushing, M. P. Walkup, B. Nicklas, W. J. Rejeski, J. P. Williams, A. Xia, B. G. Albertson, E. R. Barton, F. W. Booth, T. Caputo, M. Cicha, L. G. O. De Sousa, R. Farrar, A. L. Hevener, M. F. Hirshman, B. E. Jackson, B. G. Ke, K. S. Kramer, S. J. Lessard, N. S. Makarewicz, A. G. Marshall, P. Nigro, S. Powers, K. Ramachandran, R. S. Rector, C. Z.-T. Richards, J. Thyfault, Z. Yan, C. Zang, M. A. S. Amper, A. T. Balci, C. Chavez, M. Chikina, R. Chiu, M. A. Gritsenko, K. Guevara, J. R. Hansen, K. M. Hennig, C.-J. Hung, C. Hutchinson-Bunch, C. A. Jin, X. Liu, K. M. Maner-Smith, D. R. Mani, N. Marjanovic, M. E. Monroe, R. J. Moore, S. G. Moore, C. C. Mundorff, D. Nachun, M. D. Nestor, G. Nudelman, C. Pearce, V. A. Petyuk, H. Pincas, I. Ramos, A. (Sasha) Raskind, S. Rirak, J. M. Robbins, A. B. Rubenstein, F. Ruf-Zamojski, T. J. Sagendorf, N. Seenarine, T. Soni, K. Uppal, S. Vangeti, M. Vasoya, A. Vornholt, X. Yu, E. Zaslavsky, N. Zebarjadi, M. Bamman, B. C. Bergman, D. H. Bessesen, T. W. Buford, T. L. Chambers, P. M. Coen, D. Cooper, F. Haddad, K. Gadde, B. H. Goodpaster, M. Harris, K. M. Huffman, C. M. Jankowski, N. M. Johannsen, W. M. Kohrt, B. Lester, E. L. Melanson, K. L. Moreau, N. Musi, R. L. Newton, S. Radom-Aizik, M. E. Ramaker, T. Rankinen, B. B. Rasmussen, E. Ravussin, I. E. Schauer, R. S. Schwartz, L. M. Sparks, A. Thalacker-Mercer, S. Trappe, T. A. Trappe, E. Volpi, MoTrPAC Study Group, Primary Authors, Lead Analysts, Lead Data Generators, Analysts, Animal Study Leadership, Manuscript Writing Group Leads, Manuscript Writing Group, Senior Leadership, Co-corresponding Authors, MoTrPAC Study Group, Bioinformatics Center, Biospecimens Repository, Administrative Coordinating Center, A. Data Management and Quality Control Center, Exercise Intervention Core, NIH, Preclinical Animal Study Sites, Chemical Analysis Sites and Clinical Sites, *Nature*, 2024, **629**, 174–183.

119 Y. He, Y. Zhang, J. Lai, S. Ma, P. Yin, Z. Wu and J. Zhou, *Sci. Data*, 2025, **12**, 448.

120 M. Parstorfer, G. Poschet, K. Brüning and B. Friedmann-Bette, *Physiol. Rep.*, 2025, **13**, e70206.

121 D. C. Nieman, J. Scherr, B. Luo, M. P. Meaney, D. Dréau, W. Sha, D. A. Dew, D. A. Henson and K. L. Pappan, *PLoS One*, 2014, **9**, e113725.

122 D. C. Nieman, N. D. Gillitt, W. Sha, D. Esposito and S. Ramamoorthy, *PLoS One*, 2018, **13**, e0194843.

123 D. J. Owens and S. Bennett, *Exp. Physiol.*, 2024, **109**, 1066–1079.

124 E. L. Schymanski, J. Jeon, R. Gulde, K. Fenner, M. Ruff, H. P. Singer and J. Hollender, *Environ. Sci. Technol.*, 2014, **48**, 2097–2098.

