Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Urinary metabolomic signatures of muscle injury and recovery in elite football players

Guillermo Quintas

gquintas@leitat.org

Leitat Technological Center

Ricard Pruna

FC Barcelona

Manuel Wong

FC Barcelona

Sandra Mecho

FC Barcelona

Pau Madrero

FC Barcelona

Juan Daniel Sanjuan-Herráez

Leitat Technological Center

Gil Rodas

Leitat Technological Center

Research Article

Keywords: muscle injury, metabolomics, pathway analysis, multivariate analysis

Posted Date: June 26th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10089228/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at Metabolomics on September 23rd, 2026.
See the published version at https://doi.org/10.1007/s11306-026-02530-5.

Abstract

Introduction.

Muscle injuries are a leading cause of time-loss in elite football, highlighting the need for objective biomarkers capable of reflecting the athlete's physiological state. Monitoring strategies rely predominantly on external load metrics and may not fully capture internal physiological responses to training and competition. Urinary metabolomics offers a non-invasive approach to characterize injury related physiological processes.

Objectives.

To identify urinary metabolomic signatures associated with muscle injury in elite football players and to explore their potential utility for monitoring injury-related physiological stress and recovery.

Methods.

This observational longitudinal study included 287 urine samples collected from 121 elite male football players across two consecutive seasons. Samples were clinically classified as muscle injury (n = 30) or control (n = 257). Targeted UPLC-MS quantified amino acids and tryptophan-related metabolites. Univariate analysis, pathway over-representation analysis, and PLS-DA were applied to identify injury-associated metabolic patterns.

Results.

Muscle injury was associated with alterations in amino acid turnover, purine metabolism, energetic stress, and tryptophan catabolism. Pathway analysis identified arginine and proline metabolism as the most significantly enriched. PLS-DA identified an injury-associated metabolic signature with a cross-validated AUROC of 0.88 (sensitivity 73%, specificity 88%; $p < 0.003$). Key discriminant metabolites included β-aminoisobutyric acid, hypoxanthine, xanthurenic acid, β-alanine, and alanine. Longitudinal analyses showed progressive normalization of metabolomic profiles during rehabilitation.

Conclusions.

Urinary metabolomics identifies coordinated metabolic alterations associated with muscle injury in elite football players, supporting its potential as a complementary tool to external load monitoring and for individualized injury and recovery assessment.

Introduction

The design of training programs critically depends on the effective management of external loads (e.g., running distance, speed) which are typically guided by predefined objectives or adjusted according to feedback from internal physiological responses. These include physiological parameters (e.g., heart rate recovery and blood lactate levels) and subjective indicators such as perceived exertion, fatigue, and overall wellness. However, current athlete monitoring strategies rely predominantly on external load metrics derived from electronic performance tracking systems (EPTS) [1], which do not fully capture the individual physiological stress and biological cost of training and competition. This is indeed one of the major challenges in sports medicine because the biological response to a given training load is highly individual and influenced by a complex network of metabolic, inflammatory, and recovery-related processes. Metabolomics has emerged as a promising tool to bridge this gap by providing a snapshot of the athlete systemic biological status. Complementary to conventional monitoring systems (EPTS, fatigue, etc), metabolomic profiling could reveal subtle biochemical alterations associated with excessive muscle stress and tissue damage, even before clinical manifestations, offering a potentially transformative approach for injury surveillance in professional athletes[2–6].

Despite this careful personalized monitoring, muscle injuries linked to suboptimal physiological adaptations to external loads are one of the main causes of time-loss and performance impairment in elite football. A better characterization of the internal load and physiological stress experienced by the players could complement external load surrogates, thereby improving the personalization of training programs and optimizing adaptation and recovery to training and match demands [7,8].

The metabolic demands of contracting skeletal muscles during physical activity trigger adaptations across genomic, proteomic, and metabolic levels within a complex signaling network [9,10]. Metabolomics provides a downstream functional readout of physiological adaptations integrating genomic, transcriptomic, proteomic, and environmental influences. As metabolites represent the final products of cellular regulatory processes, metabolomic profiling may offer a sensitive approach for characterizing the biological responses associated with training load, recovery, and tissue injury. In this context, metabolomics has gained interest in the field of sports physiology in recent years, as it might provide a particularly informative perspective.

Previous studies of the effects of physical exercise on metabolism across both long-term adaptations and immediate, short-term responses have reported changes key metabolic processes including ATP metabolism, glycolysis, beta-oxidation of free fatty acids and ketone bodies, as well as the association between long-term physical exercise and the upregulation of antioxidant systems [11],[12],[13],[14],[15],[16]. Furthermore, the analysis of the association between urinary metabolomic profiles and EPTS-derived metrics identified alterations in steroid hormones, hypoxanthine-related metabolites, tyrosine, tryptophan, and riboflavin pathways as well as in amino acids, and intermediates involved in phenylalanine metabolism [17].

Acute metabolic adaptations triggered by strenuous physical exertion involve a deep shift in skeletal muscle bioenergetics. This process is predominantly characterized by an acceleration of the ATP-

phosphocreatine and anaerobic glycolytic pathways, alongside concurrent elevations in blood lactate production and purine degradation cascades [18]. In adolescent male football players, exercise-induced fatigue has been demonstrated to alter interconnected biochemical networks, including nitrogen, tyrosine, and glycerophospholipid pathways, as well as the tricarboxylic acid (TCA) cycle and glycine-serine-threonine metabolism [19]. Moreover, variations in cumulative training load and the implementation of distinct training regimens significantly modulate the systemic metabolome. For instance, an evaluation of metabolic variation in samples collected from football players both immediately after and 20 hours following competitive matches revealed distinct metabolomic changes correlated with rating of perceived exertion (RPE) scores. These variations were localized to markers reflecting localized cellular strain (such as creatinine, creatine, and glycine) and metabolic energy flux (including pyruvate, formate, 1,3-dihydroxyacetone, and creatine) [20]. Taken together, these collective insights underscore the utility of monitoring metabolomic changes as a suitable strategy for discovering novel biomarkers indicative of workload adaptation and physiological stress during training [21].

Previous studies have mainly relied on univariate statistical analysis of the association between external load with metabolic changes, and between changes in metabolic levels and pathways [22,23]. However, metabolites are often co-regulated within biological networks and univariate statistics might not capture multivariate responses. To address this limitation, multivariate techniques such as Partial Least Squares (PLS) [24,25] has been proposed to identify biological processes. The identification of specific subsets of metabolites or pathways of interest could be used for advancing targeted metabolomics methods required for facilitating robust validation in clinical settings, improving the feasibility of clinical tests, and accelerating their translation into routine practice to support personalized monitoring and training programs.

We hypothesized that urinary metabolomic profiles may capture the biological internal load associated with muscle injury and tissue healing, providing complementary information to structural imaging and conventional athlete monitoring systems. This study aimed to evaluate urinary metabolomic profiling as a non-invasive tool for monitoring internal physiological load in elite football players and to identify metabolic signatures associated with muscle injury. In addition, we explored the feasibility of developing a metabolite-based non-invasive score capable of reflecting injury-related physiological stress and recovery dynamics during rehabilitation.

Materials and methods

Subjects and experimental approach

FC Barcelona carried out this observational, longitudinal study following relevant guidelines and regulations as previously described[17]. Institutional board approval for the study was obtained from the Ethics Commission of the Consell Català de l'Esport (Code 03/2019/CEICEGC, Generalitat de Catalunya, Barcelona, Spain). Urine samples and lesion status were collected from volunteers that gave written informed consent to participate in the study. To ensure confidentiality, all performance and biochemical

data were pseudo-anonymized. The sample set included **287** first morning urine samples from **121** male professional football players collected throughout **2** consecutive seasons. Although the study population was imbalanced (30 injured athletes versus 257 controls), the larger control cohort was intentionally retained to capture the substantial biological and environmental variability inherent to elite players. Urinary metabolomic profiles are influenced by physiological and behavioral factors and so, a broad representation of non-injured athletes was considered necessary to adequately characterize the reference metabolic space against which injury-associated alterations could be identified. Urine samples were collected, aliquoted and stored at -80°C until analysis. Nutritional records were not available in this study and will be incorporated in future studies.

Muscle injury assessment

Muscle injuries were defined and recorded according to the UEFA consensus recommendations for injury surveillance in professional football. An injury was considered any physical complaint sustained during training or match play that resulted in the player being unable to fully participate in future football activities (time-loss definition)[26]. All diagnoses were made by the same medical staff throughout the study period to ensure consistency. Clinical evaluation was supported by imaging techniques, including ultrasound and/or magnetic resonance imaging (MRI), when clinically indicated. Diagnosis, prognosis, treatment, rehabilitation, and return-to-play decisions followed the FC Barcelona muscle injury management guidelines [27]. For players sustaining muscle injuries, MRI examinations were performed within 24 hours after injury and subsequently repeated during the early (2–3 weeks), intermediate (4–6 weeks), and late (6–8 weeks) phases of healing to assess connective tissue repair, scar maturation, restoration of tissue tension, and edema evolution. MRI interpretation and healing assessment followed the imaging criteria and muscle healing framework previously described [28].

MRI examinations were performed using a 3-T scanner (Vantage Galan; Canon Medical Systems, Tochigi, Japan). The follow-up protocol for muscle injury healing included axial PDFS, coronal PDFS, and axial T1-weighted sequences. Imaging parameters were as follows: axial PDFS (TR/TE, 2900/33 ms; FOV, 23 × 23 cm; matrix, 352 × 352; slice thickness, 3.5 mm), coronal PDFS (TR/TE, 2700/44 ms; FOV, 25 × 25 cm; matrix, 256 × 304; slice thickness, 2.5 mm; gap, 0.4 mm), and axial T1-weighted imaging (TR/TE, 740/10 ms; FOV, 22 × 22 cm; matrix, 352 × 416; slice thickness, 3.5 mm).

Characteristics of muscle injuries

A total of 30 urine samples were collected from 10 players during periods associated with clinically diagnosed muscle injuries. Injury diagnosis and follow-up were performed according to the institutional muscle injury protocol. Muscle injuries involved different anatomical locations and severities, representing the spectrum of lesions commonly observed in elite professional football including grade 2 distal biceps femoris strain (13 samples), general hamstring strain (multiple hamstring muscles) (5 samples), grade 1–2 semimembranosus/semitendinosus/gracilis strain (8 samples), grade 2 rectus

femoris strain at septum (1 sample), grade 1 rectus femoris strain at septum (1 sample), grade 2 proximal biceps femoris strain (1 sample), and quadriceps strain (1 sample).

Analysis of amino acids

Samples were thawed, vortexed for 15 s and centrifuged at 10000 × *g* for 10 min at 4°C. 50 µL of samples were mixed with 300 µL H_2O. Then, samples were derivatized (AccQTag Ultra, Waters) using the following procedure: 10 µL of diluted sample are transferred to a 1.5 mL Eppendorf vial and mixed with 70 µL borate buffer. After mixing (vortex, 15 s), 20 µL of 6-aminoquinolyl-N-hydroxysuccinimidyl carbamate are added. The sample was vortexed (15 s) and incubated at 55°C for 10 min. Then, the sample is transferred to a 96-well plate for UPLC-MSMS analysis within 24 h. Quantitative analysis of the following amino acids by UPLC-MSMS was carried out following derivatization (AccQ-Tag). MRM transitions are summarized in **Supplementary Material SM1**.

Analysis of tryptophan-related metabolites

Samples were thawed at room temperature, vortexed for 15 s and centrifuged at 10000 × *g* for 15 min at 4°C. 50 µL of samples were placed into 1.5 mL Eppendorf tubes and mixed with 50 µL of an internal standard solution containing hydroxytryptophan-D_4, L-Kynurenine-D_4, indole-D_5-3-acetamide, 4-chloro-kynurenine-$^{13}C_2$,^{15}N, 6-hydroxymelatonin-D_4, kynurenic acid-D_5, PAGN-D_5, serotonin-D_4, tryptamine-D_4, tryptophan-D_5, xanthurenic acid-D_4 and phenylalanine-D_5 (900 nM each). Samples are mixed (vortex, 15 s) and transferred to a 96-well plate for UPLC-MSMS analysis. The quantification of the set of metabolites including tryptophan and phenylalanine metabolites was carried out on an Acquity-Xevo TQS (Waters, Milford, USA) system equipped with an electrospray ionization source using a previously developed method [3]. Samples were analysed using an Acquity HSS T3 C_{18} (100 × 2.1 mm, 1.8 µm) column. Mobile phases were H2O (0.1% v/v HCOOH) (A) and (0.1% v/v HCOOH) CH3CN (B). The gradient elution was as follows: phase B was held 2% from 0 to 0.5 min, then increased linearly to 45% over the following 5 min. Then phase B was increased to 90% in 0.2 min followed by a fast return to initial conditions between 5.7 and 6 min, which were held for 1.5 min for column re-equilibration. Injection volume, flow rate and column temperature were set at 3 µL, 550 µL/min and 55°C, respectively. Autosampler temperature was set at 6°C during sample analysis. Electrospray ionization was carried out using the following conditions: capillary 2.9 kV, cone 25 V, source temperature 120°C, desolvation temperature 395°C, N_2 cone and desolvation gas flow rates were 150 and 800 L/h, respectively.

Statistical analysis

Univariate and multivariate analysis was carried out in MATLAB (Mathworks Inc., Natick, MA, USA). Principal Component Analysis (PCA) and PLS regression models were developed using PLS Toolbox 9.5 (Eigenvector Research Inc., Wenatchee, WA, USA) and MATLAB 2023b scripts. The number of latent

variables of the PLS models was chosen based on the classification error estimated using cross validation (CV). Given the limited number of injury samples available (n = 30), leave-one-out cross-validation (LOOCV) was used to maximize the amount of information available for model training while still providing an internal estimate of predictive performance. Under LOOCV, each sample is excluded once and predicted using a model developed from all remaining samples, thereby minimizing the loss of training data associated with more restrictive partitioning schemes such as leave-one player-out or random k-fold CV. This consideration was particularly relevant in the present study because muscle injury samples represented a relatively small fraction of the dataset. The primary objective of the multivariate analysis was to identify metabolomic signatures associated with the injury state rather than to evaluate generalization to entirely unseen athletes. The statistical significance of figures of merit estimated by CV was assessed using permutation testing. The permutation test involved the random permutation of the vector of labels (control, injury), the subsequent development of a PLS-DA model using the permuted class labels, and the estimation of the mean classification error of CV (CV_{perm}) of the PLS model [29]. This step was repeated 300 times, to obtain a null distribution of CV_{perm} performance values that was used to compute the p-value for the performance estimation obtained using the true class labels (CV_{init}), as the fraction of permuted statistics $CV_{perm} < CV_{init}$.. Permutation testing was performed by randomly permuting sample class labels. Although multiple samples were available from some athletes, player-level permutation strategies were not considered appropriate because muscle injury represented a transient condition and most players contributed only non-injury samples, which would substantially restrict the number of valid subject-level permutations and result in an uninformative null distribution.

Backward stepwise elimination was used as a feature selection algorithm by excluding iteratively individual features one by one to maximize model performance. First, the search algorithm builds a PLS-DA model where the number of latent variables is selected based on the estimated classification error as described before by LOOCV. Then, the Variable Importance in the Projection (VIP) score vector is calculated, and the variable showing the lowest VIP score is identified and excluded from the selected variable subset. The process is iterated until the size of the selected variable subset is lower than the maximum number of latent variables evaluated (3 in this study). Metabolic functional analysis was carried out using the *mummichog* algorithm[30,31] available in MetaboAnalyst 6.0[32].

Results and discussion

Overview of the data set

A total of 287 urine samples collected from 121 elite male football players across two consecutive seasons were analyzed using targeted UPLC-MS metabolomics for the quantification of 56 metabolic ratios after normalization by creatinine concentration and 15 after normalization by Trp concentrations. Samples were categorized into two clinical conditions: muscle injury (n = 30) and non-injured controls (n = 257).

Prior to supervised statistical analysis, principal component analysis (PCA) was performed to assess the global structure of the dataset, identify potential analytical outliers, and evaluate the major sources of variance. The PCA model showed the overlap between samples obtained during muscle injury periods and samples collected during non-injured control conditions (Fig. 1). No clustering according to injury status was observed in the two first principal components, indicating that muscle injury was not among the main sources of variance within the metabolic dataset. No analytical outliers were detected, supporting the robustness of the sample collection and analytical workflows. In addition, the distribution of samples according to season or year of collection did not reveal systematic batch-related separation, suggesting that inter-season variability had a limited influence on the data structure.

From a systems perspective, the urinary metabolome reflects the integrated outcome of external load exposure, internal physiological strain, and injury-recovery status, among other factors. The observed overlap between injured and non-injured states, indicating that muscle injury responses do not dominate global metabolic variance is consistent with training load theory, where injury-related perturbations are expected to be subtle compared with the broader physiological variability observed in football players.

Discriminant analysis

To identify metabolites associated with muscle injury, univariate analyses combining Welch's test, fold-change assessment, and Hedge's *g* effect size estimation were performed using winsorized metabolite data (Fig. 2). Metabolites showing significant differences were subsequently subjected to pathway over-representation analysis (ORA) using the KEGG database.

Several metabolites showed statistically significant associations with muscle injury, involving coordinated alterations in pathways related to energetic metabolism, amino acid turnover, tryptophan metabolism, and cellular stress responses. In addition to statistical significance, effect sizes were evaluated using Hedge's *g* to quantify the magnitude of group differences (Fig. 2, center panel). Most significant metabolites displayed moderate-to-large effect sizes, supporting the biological relevance of the observed alterations. The largest positive effect sizes were observed for metabolites related to tryptophan metabolism, including xanthurenic acid, kynurenic acid, and tryptamine/tryptophan ratio, as well as S-adenosylmethionine, indicating substantially higher levels in injured players. Conversely, aminophenol, proline, ornithine, and lysine exhibited negative effect sizes, reflecting lower concentrations in the injury group. The overall concordance between Welch's test significance, fold-change analysis, and Hedge's *g* increases the confidence that these metabolites represent meaningful injury-associated metabolic alterations rather than isolated statistical findings.

A metabolic change associated with muscle injury was the alteration of metabolites linked to purine degradation, particularly hypoxanthine, together with changes in methylation-related intermediate such as S-adenosylmethionine. This pattern could indicate increased ATP turnover and energetic stress at the cellular level. Purine catabolism products such as hypoxanthine are well-established markers of energetic stress, reflecting the breakdown of ATP during periods of increased metabolic demand. In the

context of muscle injury, these changes could reflect the joint effect of localized energy depletion within damaged muscle fibers and systemic metabolic adjustments required to sustain tissue repair processes. As mentioned before, a second major component of the injury-associated metabolic changes involved alterations in amino acids and amino acid–related metabolites, including alanine, aminoadipic acid, glutamic acid, lysine, ornithine, proline, taurine, β-alanine, and phenylacetylglutamine. Again, rather than representing independent metabolic perturbations, collectively these changes could be linked to coordinated protein turnover and tissue remodeling. Several of these metabolites are directly implicated in skeletal muscle structural remodeling. For instance, proline and ornithine are closely linked to collagen synthesis and connective tissue repair. The elevation of BAIBA further could be linked to changes in the activity of metabolic pathways associated with muscle stress adaptation and mitochondrial remodeling. A third component involved significant alterations in tryptophan metabolism, particularly through the kynurenine pathway and related indole derivatives. Metabolites belonging to the kynurenine metabolism displayed both significant p-values and among the largest effect sizes, suggesting that perturbation of this pathway represents one of the most relevant metabolic responses associated with muscle injury. Increased levels of kynurenic acid and xanthurenic acid, together with changes in indole-3-acetic acid, indole-3-acetamide, and tryptamine, indicate a broad reprogramming of tryptophan metabolism during muscle injury. The kynurenine pathway is strongly regulated by immune signaling and inflammatory mediators and has been widely implicated in exercise-induced physiological adaptation and inflammatory regulation [33,34]. Furthermore, indolic metabolites likely reflect interactions between host metabolism and gut microbial activity.

Although these metabolic changes can be categorized into distinct biochemical components, their simultaneous change suggests a coordinated systemic adaptation to muscle injury. Thus, this integrated response likely reflects the combined demands of tissue damage repair, inflammatory regulation, and metabolic demands in elite players exposed to high external and internal loads. These results support the potential utility of urinary metabolomics as a non-invasive tool for monitoring physiological responses to injury.

The metabolites showing significant Welch's test results and moderate-to-large Hedge's *g* values were subsequently used for pathway over-representation analysis (ORA) (see Fig. 2). This analysis showed a significant enrichment of amino acid metabolic pathways, particularly arginine and proline metabolism, indicating that the individual metabolite alterations converge on processes that might be associated with tissue repair and remodeling.

Multivariate discriminant analysis

PLS-DA was performed to identify coordinated metabolic patterns associated with muscle injury. Despite the significant imbalance between injury and control samples, the PLS-DA model showed statistically significant classification performance (see Fig. 3). Cross-validated predictions correctly classified 73% of injury samples and 88% of non-injured samples, yielding an overall AUROC(CV) of 0.88. The positive predictive value was 41% (95% CI 32–50%), and the negative predictive value 97% (95% CI 94–98%). The

F1-score and the Matthews correlation coefficient (MCC) were 0.52 and 0.48, respectively. The F1-score indicated a robust balance between injury detection and false positive mitigation under highly imbalanced conditions (~ 1:9 ratio in this study), while the MCC value confirms that the model's predictive power is significantly better than random guessing and resilient to class prevalence. Furthermore, permutation testing supported the statistical significance of the model (AUROC-CV p-value < 0.003).

To ensure that the LOOCV approach did not introduce bias or data leakage due to the multi-sample contribution of individual athletes, a sensitivity analysis was performed by iteratively excluding entire player trajectories (leave-one-player-out, LOPO-CV). A substantial reduction in model performance was observed (AUROC-CV = 0.72, F1-score = 0.32, MCC = 0.22) which was attributed to the effect of LOPO in a highly imbalanced set (30 injury vs. 257 control samples) in which the 30 injury samples originated from a pool of 10 distinct players. This confirmed that sample-level LOOCV was required to maintain a representative calibration space capable of handling inter-individual metabolic differences among football players.

The regression vector and variable importance in projection (VIP) analysis identified BAIBA, hypoxanthine, xanthurenic acid, β-alanine, and alanine among the metabolites contributing most strongly to sample discrimination (Fig. 3). The metabolites identified through multivariate modeling were largely consistent with the results obtained from univariate analyses and pathway enrichment approaches, supporting the robustness and biological coherence of the identified metabolic signatures.

The predicted probability distributions obtained by CV showed partial overlap between injury and control samples. Nevertheless, the statistically significant classification performance indicates that coordinated metabolic information contained within urinary metabolomic profiles can capture physiologically relevant signatures associated with muscle injury.

Feature selection

To identify a reduced subset of metabolites with maximal predictive relevance, backward stepwise feature selection was applied to the PLS-DA model. Variables with lower VIP scores were progressively removed values while monitoring model performance during CV. Feature selection resulted in a restricted subset of metabolites that preserved classification performance while substantially simplifying the model structure (Fig. 4). The selected metabolic set included metabolites associated with tryptophan metabolism, purine metabolism, amino acid turnover, and energetic stress pathways.

The identification of a limited panel of discriminant metabolites is particularly relevant from a translational perspective. Compared to large-scale quantitative or untargeted metabolomics, reduced targeted panels may provide improved analytical robustness and facilitate the development and validation of quantitative assays for routine athlete monitoring. Using the selected subset of metabolites, a metabolite-based score reflecting an injury-associated metabolic signature was constructed. Although

validation in independent cohorts would be required, these findings support the feasibility of developing targeted metabolomic tools for individualized athlete monitoring and recovery assessment.

Longitudinal profiles

Longitudinal metabolomic profiles were available for several players with serial urine samples collected throughout injury and recovery periods. Overall, a visual exploratory analysis of the longitudinal trajectories of the metabolic injury scores showed the progressive normalization of the metabolomic profiles during rehabilitation up to the return-to-play (RTP) milestone. Metabolic profiles altered during the acute phase of injury progressively approached baseline control values as physiological recovery progressed (Fig. 5). This temporal behavior supports the hypothesis that urinary metabolomic profiles dynamically reflect physiological recovery processes associated with tissue repair or its metabolic adaptation. Although inter-individual variability was observed, the general longitudinal trends were biologically coherent across players, including those who sustained no muscle injuries throughout the study period, demonstrating the baseline stability of the score.

The practical application of this joint monitoring framework is shown by the serial MRI follow-up (axial and coronal Proton Density Fat-Suppressed [PDFS] sequences) of a 24-year-old professional football player who sustained a high-speed running injury involving the biceps femoris (Fig. 6). MRI obtained within 24 hours of injury demonstrates a complete tear of the proximal common hamstring aponeurosis (arrow) located 10.5 cm distal to its ischial tuberosity origin, with distal retraction, loss of connective tissue tension, interstitial edema, and intermuscular edema (MLG-R grade IPp3r0 BAMIC grade 3C) (Fig. 6A). Three weeks after injury, the tendon gap (arrow) is bridged by a hypertrophic reparative scar (phase 2a), with restoration of connective tissue tension and reduction of interstitial and intermuscular edema (Fig. 6B). Six weeks after injury, the scar had progressed to a phase 2b–3a appearance (arrow), characterized by a more uniform hypointense peripheral rim and decreasing central hyperintensity, reflecting ongoing scar maturation. Interstitial edema has resolved, with only mild adaptive edema and minimal residual intermuscular edema remaining (Fig. 6C). At eight weeks after injury, the scar was predominantly hypointense (arrow), consistent with advanced remodeling and connective tissue maturation, with only minimal adaptive edema persisting (Fig. 6D).

Analysing the joint evolution of structural lesions alongside metabolic profiles could be highly important for elite sports medicine. While conventional clinical milestones rely heavily on the macro-structural restoration of muscle architecture seen on MRI (as shown in Fig. 6), macro-structural adaptation might not always reflect the resolution of local energetic or systemic inflammatory stress. Thus, the mapping of metabolomic trajectories relative to structural lesion evolution could support the assessment of the recovery of the athlete's cellular homeostasis, or if persistent metabolic alterations such as elevated purine degradation or altered tryptophan catabolism remain despite normal structural imaging findings.

Conclusions and outlook

This study demonstrates that urinary metabolomic profiling can detect coordinated metabolic alterations associated with muscle injury in elite football players. While injury-related changes were subtle, the combination of univariate analysis, pathway enrichment, and multivariate classification identified biological signatures primarily involving amino acid turnover, tryptophan catabolism, and energetic stress (purine metabolism) associated with muscle injury. These findings suggest that integrated, longitudinal monitoring incorporating both the localized structural evolution of the lesion and non-invasive urinary metabolomic scores could provide an objective perspective to guide individual programming and safely manage return-to-play timelines in elite football players.

Several limitations should be considered when interpreting the results of this study. First, nutritional intake and other potential confounding factors such as acute training workload or medication which might influence the urinary metabolomic were not recorded. Second, although a large longitudinal database of urine samples was analyzed, the number of clinically confirmed muscle injury events remained limited and restricted to a single-center cohort of 10 distinct players. Third, urinary metabolomics reflects systemic metabolic responses and may therefore capture physiological processes not exclusively related to skeletal muscle injury. Finally, the observational nature of the study limits causal interpretation of the identified associations.

Future multicenter studies should focus on validating these urinary signatures in larger independent cohorts. Integrating longitudinal metabolomics with structural imaging datasets tracking lesion evolution, biochemical markers, and clinical assessments may support the analysis of the synchronous relationship between tissue remodeling and metabolic normalization. This multimodal approach has the potential to advance personalized recovery monitoring and facilitate individualized return-to-play decision-making guidelines in professional sports.

Declarations

Acknowledgements

The authors acknowledge the dedicated support from the FCB volunteers, FCB Medical Service, and Barça Innovation Hub (https://barcainnovationhub.com/).

Competing interests

The authors declare no competing interests.

Author Contribution Statement

GQ: methodology, performed research, analyzed data; RP: resources, analyzed data; MW: resources, analyzed data; SM: resources, analyzed data; PM: resources, analyzed data; JDS: analysis; GR: conceptualization, methodology, analysis, resources. All authors wrote and reviewed the manuscript.

Compliance with Ethical Standards

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. Informed consent was obtained from all individual participants included in the study.

Data Availability Statement

Due to the highly specific nature of the cohort and the inclusion of longitudinal clinical, performance, and injury-related temporal information, the dataset carries a non-negligible risk of participant re-identification even after pseudonymisation. In accordance with the conditions of the institutional ethics approval and applicable data protection regulations, the raw dataset is therefore not publicly available.

References

1. R. Akenhead, G.P. Nassis, Training Load and Player Monitoring in High-Level Football: Current Practice and Perceptions, Int. J. Sports Physiol. Perform. 11 (2016) 587–593. https://doi.org/10.1123/ijspp.2015-0331.
2. K. Khoramipour, Ø. Sandbakk, A.H. Keshteli, A.A. Gaeini, D.S. Wishart, K. Chamari, Metabolomics in Exercise and Sports: A Systematic Review, Sports Med. 52 (2022) 547–583. https://doi.org/10.1007/s40279-021-01582-y.
3. G. Rodas, E. Ferrer, X. Reche, J.D. Sanjuan-Herráez, A. McCall, G. Quintás, A targeted metabolic analysis of football players and its association to player load: Comparison between women and men profiles, Front. Physiol. 13 (2022). https://doi.org/10.3389/fphys.2022.923608.
4. J.R. González, A. Cáceres, E. Ferrer, L. Balagué-Dobón, X. Escribà-Montagut, D. Sarrat-González, G. Quintás, G. Rodas, Predicting Injuries in Elite Female Football Players With Global-Positioning-System and Multiomics Data, Int. J. Sports Physiol. Perform. 19 (2024) 661–669. https://doi.org/10.1123/ijspp.2023-0184.
5. H. Schoeny, B. Stelzer, T. Hofbauer, F. Reisenbauer, Y.E. Abiead, J. Scharhag, G. Koellensperger, Metabolome trajectories in male and female athletes, J. Sport Health Sci. 14 (2025) 101065. https://doi.org/10.1016/j.jshs.2025.101065.
6. G. Rodas, E. Ferrer, J.D. Sanjuan, G. Quintás, UPLC-MS and multivariate analysis reveal metabolic pathway adaptations to training in professional football players, Talanta 291 (2025) 127893. https://doi.org/10.1016/j.talanta.2025.127893.
7. A.L. Arcos, A. Mendez-Villanueva, R. Martínez-Santos, In-season training periodization of professional soccer players, Biol. Sport 34 (2017) 149–155. https://doi.org/10.5114/biolsport.2017.64588.
8. P.C. Bourdon, M. Cardinale, A. Murray, P. Gastin, M. Kellmann, M.C. Varley, T.J. Gabbett, A.J. Coutts, D.J. Burgess, W. Gregson, N.T. Cable, Monitoring Athlete Training Loads: Consensus Statement, Int. J. Sports Physiol. Perform. 12 (2017) S2161–S2170. https://doi.org/10.1123/IJSPP.2017-0208.

9. J.A. Hawley, M. Hargreaves, M.J. Joyner, J.R. Zierath, Integrative Biology of Exercise, Cell 159 (2014) 738–749. https://doi.org/10.1016/j.cell.2014.10.029.
10. F.M. Impellizzeri, S.M. Marcora, A.J. Coutts, Internal and External Training Load: 15 Years On, Int. J. Sports Physiol. Perform. 14 (2019) 270–273. https://doi.org/10.1123/ijspp.2018-0935.
11. J. Finaud, G. Lac, E. Filaire, Oxidative stress : relationship with exercise and training, Sports Med. Auckl. NZ 36 (2006) 327–358.
12. E.M. Gorostiaga, I. Navarro-Amézqueta, J.A.L. Calbet, Y. Hellsten, R. Cusso, M. Guerrero, C. Granados, M. González-Izal, J. Ibañez, M. Izquierdo, Energy Metabolism during Repeated Sets of Leg Press Exercise Leading to Failure or Not, PLOS ONE 7 (2012) e40621. https://doi.org/10.1371/journal.pone.0040621.
13. A. Pechlivanis, K.G. Papaioannou, G. Tsalis, P. Saraslanidis, V. Mougios, G.A. Theodoridis, Monitoring the Response of the Human Urinary Metabolome to Brief Maximal Exercise by a Combination of RP-UPLC-MS and (1)H NMR Spectroscopy, J. Proteome Res. 14 (2015) 4610–4622. https://doi.org/10.1021/acs.jproteome.5b00470.
14. R.G. Duft, A. Castro, M.P.T. Chacon-Mikahil, C.R. Cavaglieri, R.G. Duft, A. Castro, M.P.T. Chacon-Mikahil, C.R. Cavaglieri, Metabolomics and Exercise: possibilities and perspectives, Mot. Rev. Educ. Física 23 (2017). https://doi.org/10.1590/s1980-6574201700020010.
15. L.M. Heaney, K. Deighton, T. Suzuki, Non-targeted metabolomics in sport and exercise science, J. Sports Sci. 0 (2017) 1–9. https://doi.org/10.1080/02640414.2017.1305122.
16. F.A. Manaf, N.G. Lawler, J.J. Peiffer, G.L. Maker, M.C. Boyce, T.J. Fairchild, D. Broadhurst, Characterizing the plasma metabolome during and following a maximal exercise cycling test, J. Appl. Physiol. Bethesda Md 1985 125 (2018) 1193–1203. https://doi.org/10.1152/japplphysiol.00499.2018.
17. G. Quintas, X. Reche, J.D. Sanjuan-Herráez, H. Martínez, M. Herrero, X. Valle, M. Masa, G. Rodas, Urine metabolomic analysis for monitoring internal load in professional football players, Metabolomics Off. J. Metabolomic Soc. 16 (2020) 45. https://doi.org/10.1007/s11306-020-01668-0.
18. C. Enea, F. Seguin, J. Petitpas-Mulliez, N. Boildieu, N. Boisseau, N. Delpech, V. Diaz, M. Eugène, B. Dugué, (1)H NMR-based metabolomics approach for exploring urinary metabolome modifications after acute and chronic physical exercise, Anal. Bioanal. Chem. 396 (2010) 1167–1176. https://doi.org/10.1007/s00216-009-3289-4.
19. B. Cao, S. Liu, L. Yang, A. Chi, Changes of Differential Urinary Metabolites after High-Intensive Training in Teenage Football Players, BioMed Res. Int. 2020 (2020) 2073803. https://doi.org/10.1155/2020/2073803.
20. A.H. Marinho, F.A. de B. Sousa, R. de A.M.P. Vilela, P. Balikian, E. de Souza Bento, T. de Mendonça Aquino, A. Crispim, T. Ataide-Silva, G.G. de Araujo, The rating of perceived exertion is able to differentiate the post-matches metabolomic profile of elite U-20 soccer players, Eur. J. Appl. Physiol. 122 (2022) 371–382. https://doi.org/10.1007/s00421-021-04838-7.

21. C.M. Neal, A.M. Hunter, L. Brennan, A. O'Sullivan, D.L. Hamilton, G. De Vito, S.D.R. Galloway, Six weeks of a polarized training-intensity distribution leads to greater physiological and performance adaptations than a threshold model in trained cyclists, J. Appl. Physiol. Bethesda Md 1985 114 (2013) 461–471. https://doi.org/10.1152/japplphysiol.00652.2012.
22. M.A. García-Campos, J. Espinal-Enríquez, E. Hernández-Lemus, Pathway Analysis: State of the Art, Front. Physiol. 6 (2015) 383. https://doi.org/10.3389/fphys.2015.00383.
23. T.-M. Nguyen, A. Shafı, T. Nguyen, S. Draghici, Identifying significantly impacted pathways: a comprehensive review and assessment, Genome Biol. 20 (2019) 203. https://doi.org/10.1186/s13059-019-1790-4.
24. E. Saccenti, H.C.J. Hoefsloot, A.K. Smilde, J.A. Westerhuis, M.M.W.B. Hendriks, Reflections on univariate and multivariate analysis of metabolomics data, Metabolomics 10 (2014) 361–374. https://doi.org/10.1007/s11306-013-0598-6.
25. A.C. Olivieri, Analytical Advantages of Multivariate Data Processing. One, Two, Three, Infinity?, Anal. Chem. 80 (2008) 5713–5720. https://doi.org/10.1021/ac800692c.
26. C.W. Fuller, J. Ekstrand, A. Junge, T.E. Andersen, R. Bahr, J. Dvorak, M. Hägglund, P. McCrory, W.H. Meeuwisse, Consensus statement on injury definitions and data collection procedures in studies of football (soccer) injuries, Br. J. Sports Med. 40 (2006) 193–201. https://doi.org/10.1136/bjsm.2005.025270.
27. Barça Innovation Hub. Muscle Injury Guide: Prevention of and Return to Play from Muscle Injuries, (2018).
28. J. Isern-Kebschull, S. Mechó, C. Pedret, R. Pruna, X. Alomar, A. Kassarjian, A. Luna, J. Martínez, X. Tomas, G. Rodas, Muscle Healing in Sports Injuries: MRI Findings and Proposed Classification Based on a Single Institutional Experience and Clinical Observation, Radiogr. Rev. Publ. Radiol. Soc. N. Am. Inc 44 (2024) e230147. https://doi.org/10.1148/rg.230147.
29. C.M. Rubingh, S. Bijlsma, E.P.P.A. Derks, I. Bobeldijk, E.R. Verheij, S. Kochhar, A.K. Smilde, Assessing the performance of statistical validation tools for megavariate metabolomics data, Metabolomics Off. J. Metabolomic Soc. 2 (2006) 53–61. https://doi.org/10.1007/s11306-006-0022-6.
30. S. Li, Y. Park, S. Duraisingham, F.H. Strobel, N. Khan, Q.A. Soltow, D.P. Jones, B. Pulendran, Predicting Network Activity from High Throughput Metabolomics, PLOS Comput. Biol. 9 (2013) e1003123. https://doi.org/10.1371/journal.pcbi.1003123.
31. Z. Pang, G. Zhou, J. Ewald, L. Chang, O. Hacariz, N. Basu, J. Xia, Using MetaboAnalyst 5.0 for LC–HRMS spectra processing, multi-omics integration and covariate adjustment of global metabolomics data, Nat. Protoc. 17 (2022) 1735–1761. https://doi.org/10.1038/s41596-022-00710-w.
32. Z. Pang, Y. Lu, G. Zhou, F. Hui, L. Xu, C. Viau, A.F. Spigelman, P.E. MacDonald, D.S. Wishart, S. Li, J. Xia, MetaboAnalyst 6.0: towards a unified platform for metabolomics data processing, analysis and interpretation, Nucleic Acids Res. (2024) gkae253. https://doi.org/10.1093/nar/gkae253.

33. I. Cervenka, L.Z. Agudelo, J.L. Ruas, Kynurenines: Tryptophan's metabolites in exercise, inflammation, and mental health, Science 357 (2017) eaaf9794. https://doi.org/10.1126/science.aaf9794.

34. M.V. dos S. Rangel, K.G. Lopes, X. Qin, J.P. Borges, Exercise-induced adaptations in the kynurenine pathway: implications for health and disease management, Front. Sports Act. Living 7 (2025) 1535152. https://doi.org/10.3389/fspor.2025.1535152.

Figures

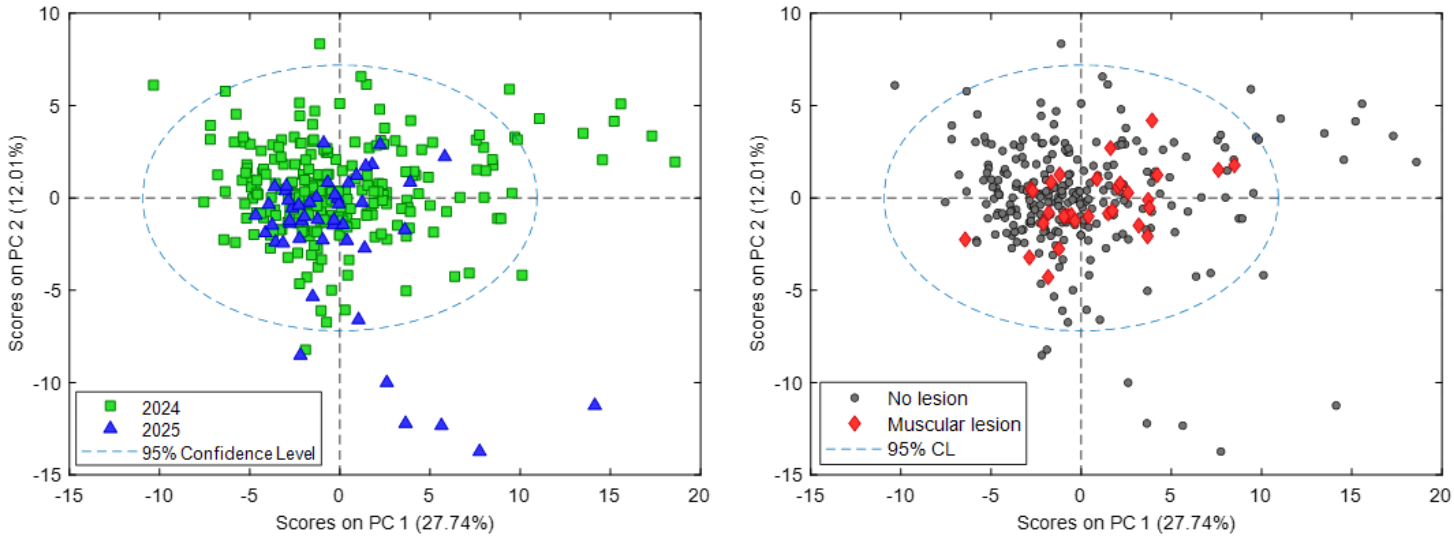


Figure 1

PCA of the full data set. Scores plot obtained from PCA analysis of urinary metabolomic profiles collected from elite football players during muscle injury and non-injured periods. Each point represents an individual urine sample.

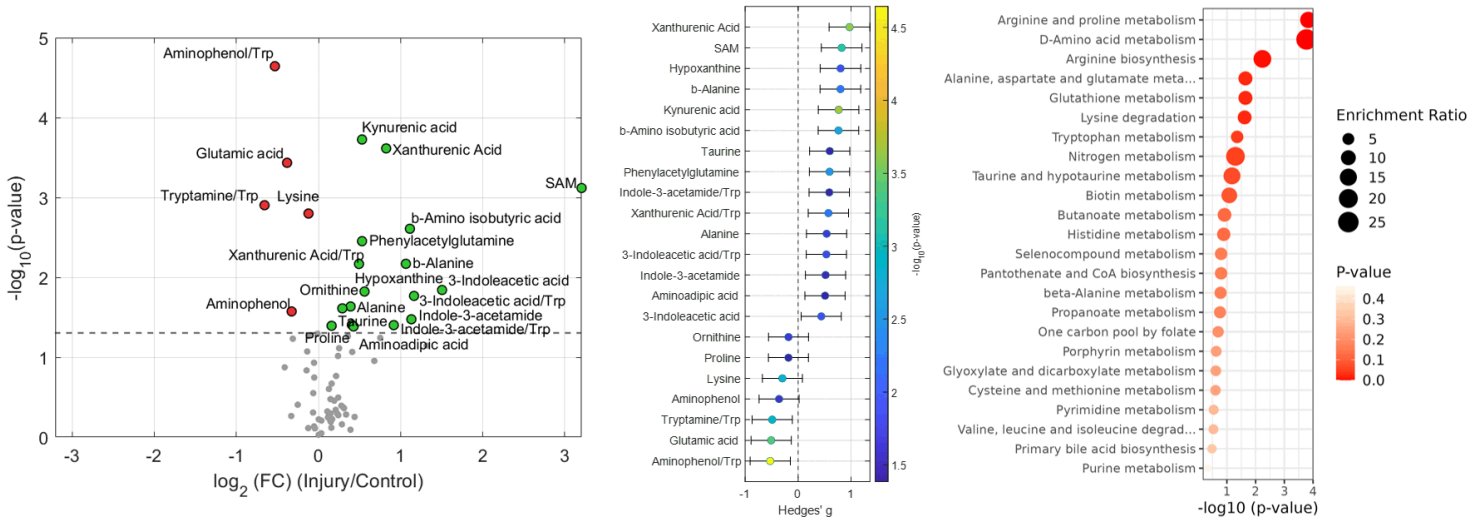


Figure 2

Differential urinary metabolites associated with muscle injury. Left) volcano plot showing differential metabolite abundance between injury and control groups. The x-axis represents $\log_2$(Fold change) (FC; Injury/Control), and the y-axis represents -$\log_{10}$(p-value Welch's test). Each point corresponds to an individual metabolite. Metabolites significantly increased in the injury group are shown in green, while significantly decreased metabolites are shown in red; non-significant metabolites are shown in gray. Center) Forest plot showing Hedge's *g* effect sizes and confidence intervals for metabolites with $p < 0.05$. Positive values indicate higher abundance in the injury group, whereas negative values indicate lower abundance relative to controls. Larger absolute Hedge's *g* values reflect greater separation between groups; Right) KEGG pathway over-representation analysis (ORA) of significant metabolites. Dot size represents enrichment ratio and color represents pathway significance (p-value).

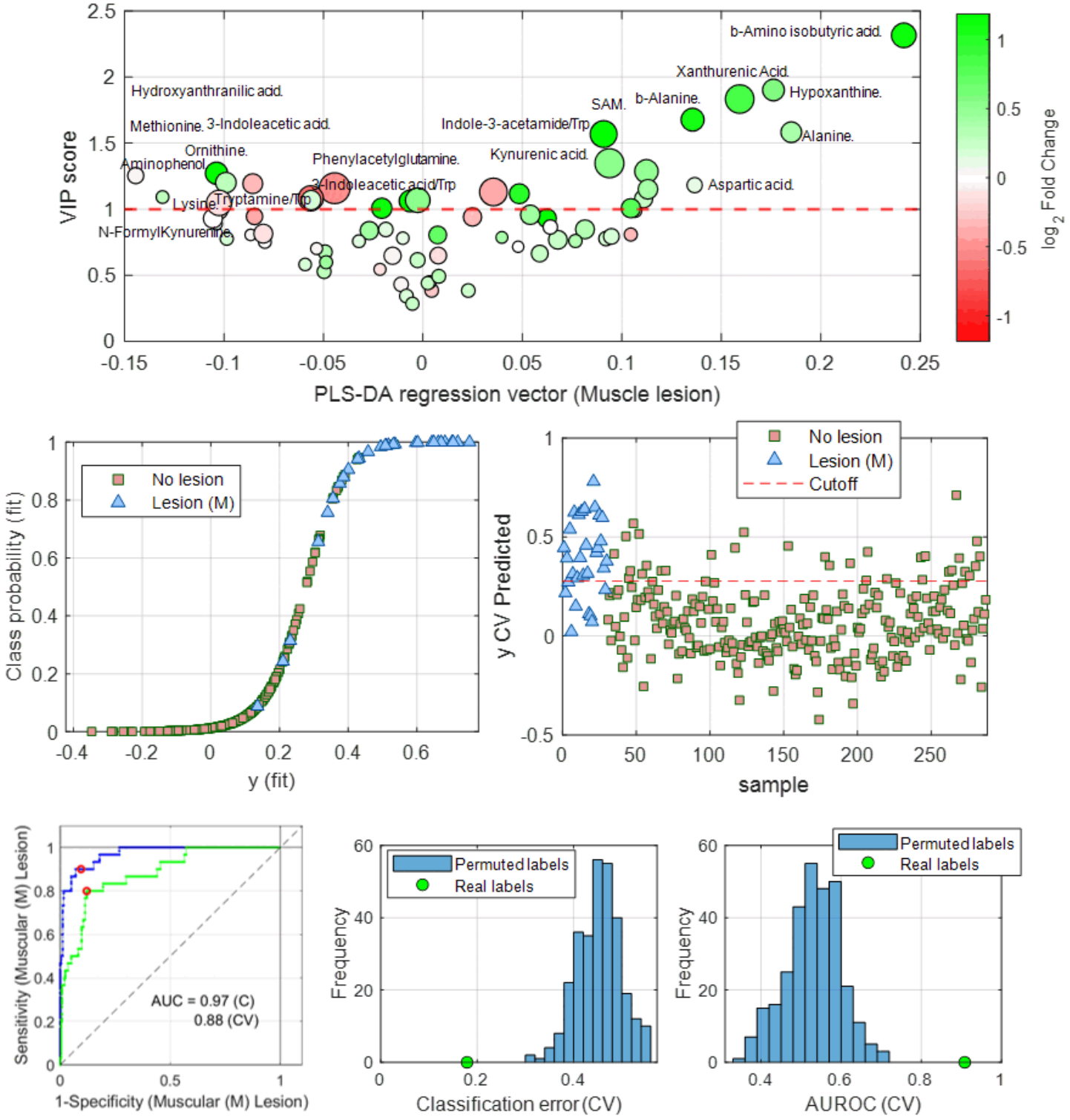


Figure 3

PLS-DA model for discrimination between muscle injury and non-injured control samples. Left) Cross-validated predicted class probabilities obtained from the PLS-DA model. Right) Variable importance in projection (VIP) scores and regression coefficients for the metabolites contributing most strongly to class discrimination.

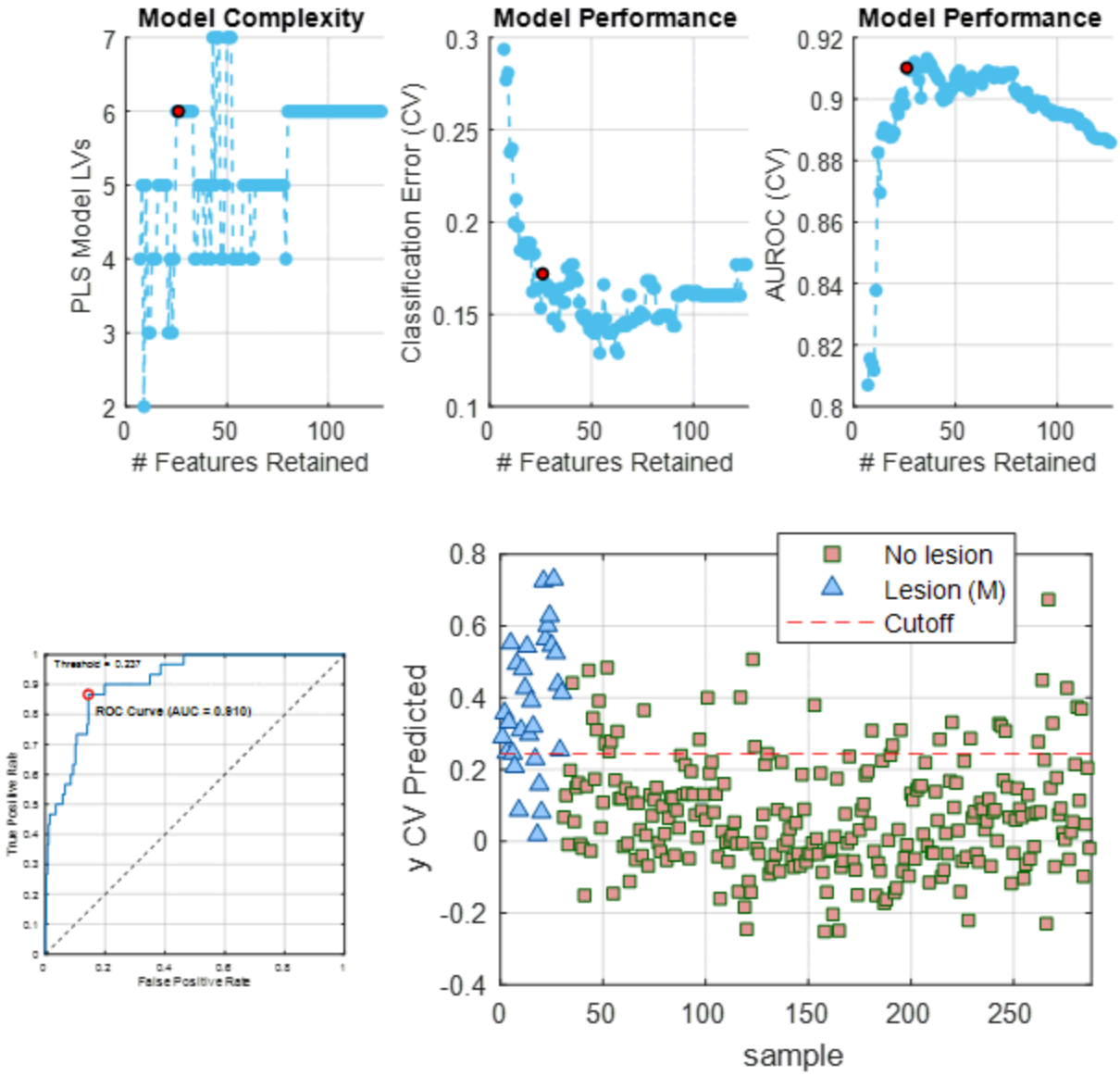


Figure 4

Backward stepwise feature selection and injury-associated metabolic score. Left) Classification performance of the PLS-DA model as a function of the number of retained metabolites during backward feature elimination. Right) Distribution of the metabolite-based injury score in injury and non-injured control samples.

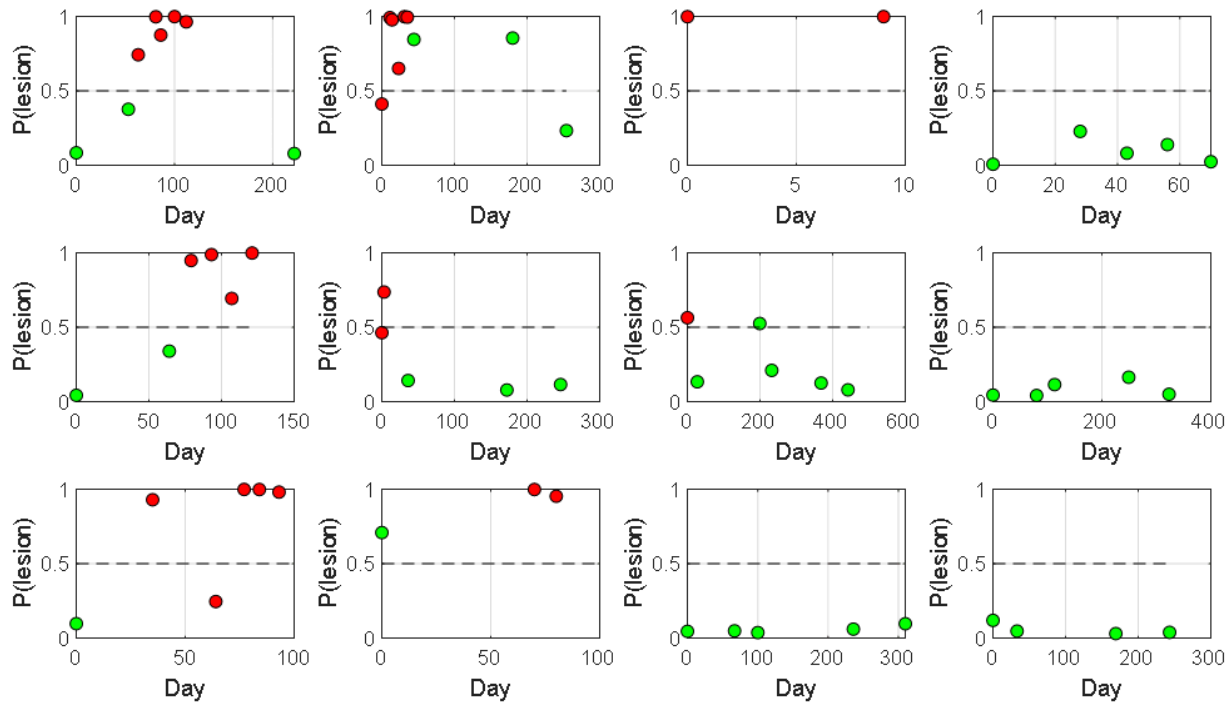


Figure 5

Representative longitudinal metabolomic trajectories during injury and recovery periods in elite football players. Metabolite profiles altered during acute injury progressively approached control values throughout rehabilitation and return-to-play phases, supporting the dynamic association between urinary metabolomic signatures and physiological recovery processes.

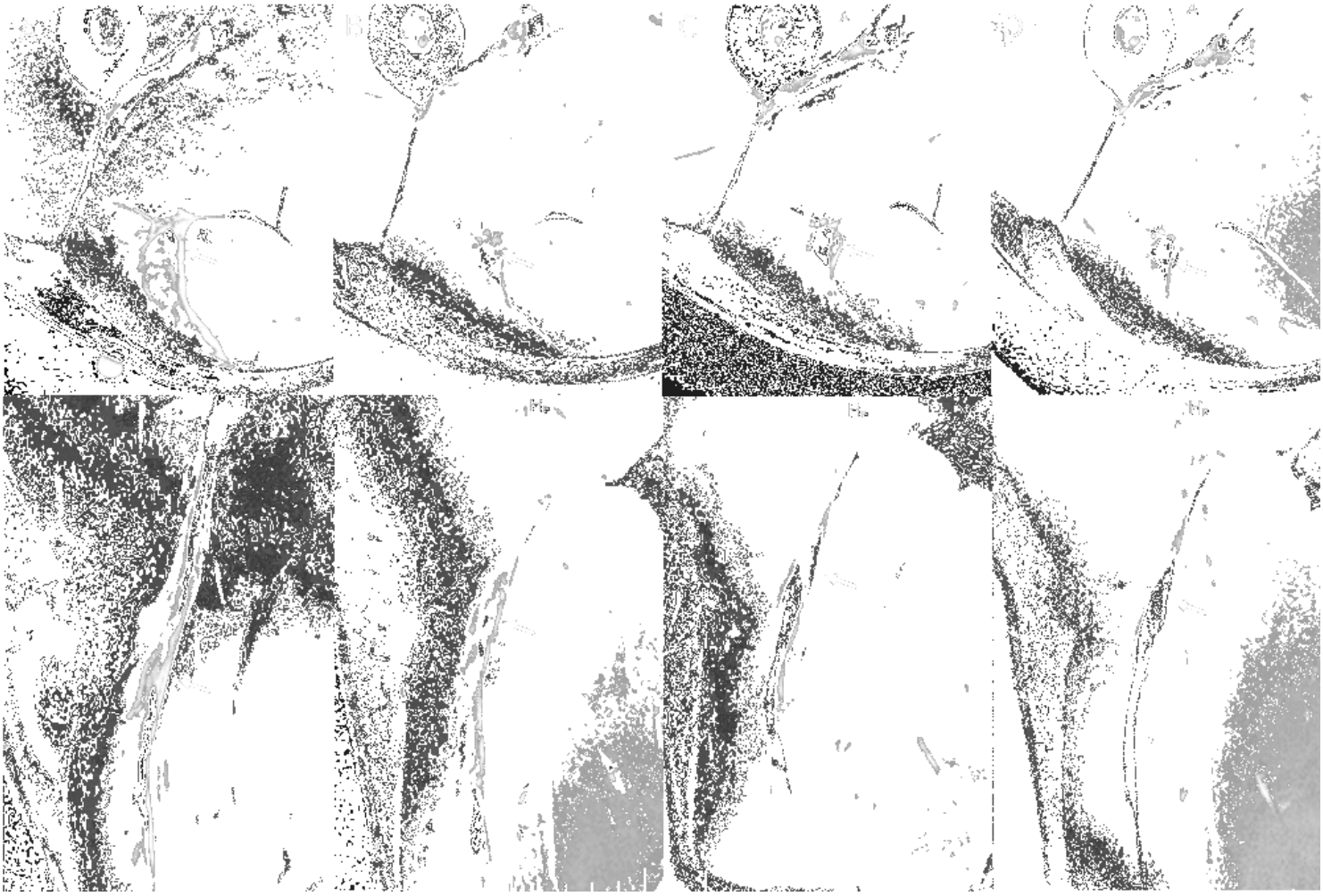


Figure 6

Longitudinal MRI assessment of muscle healing following a biceps femoris injury. Serial axial fat-suppressed T2-weighted MRI examinations acquired from the acute phase of injury (left) to return-to-play (right). Images demonstrate the progressive resolution of intramuscular edema, reduction of the connective tissue gap, restoration of muscle architecture, and maturation of the repair tissue over time. MRI follow-up was performed every 2–3 weeks as part of the FC Barcelona muscle injury management protocol. These imaging findings illustrate the dynamic biological healing process and were used in parallel with urinary metabolomic monitoring to investigate the relationship between tissue repair and systemic metabolic adaptations during recovery.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryMaterial.docx