

Northumbria Research Link

Citation: Stander, Zinandré, Luies, Laneke, Mienie, Japie, van Reenen, Mari, Howatson, Glyn, Keane, Karen, Clifford, Tom, Stevenson, Emma J. and Loots, Du Toit (2020) The unaided recovery of marathon-induced serum metabolome alterations: Post-marathon metabolic recovery. *Scientific Reports*, 10 (1). p. 11060. ISSN 2045-2322

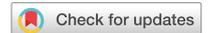
Published by: Nature Publishing

URL: <https://doi.org/10.1038/s41598-020-67884-9> <<https://doi.org/10.1038/s41598-020-67884-9>>

This version was downloaded from Northumbria Research Link:
<http://nrl.northumbria.ac.uk/id/eprint/43499/>

Northumbria University has developed Northumbria Research Link (NRL) to enable users to access the University's research output. Copyright © and moral rights for items on NRL are retained by the individual author(s) and/or other copyright owners. Single copies of full items can be reproduced, displayed or performed, and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided the authors, title and full bibliographic details are given, as well as a hyperlink and/or URL to the original metadata page. The content must not be changed in any way. Full items must not be sold commercially in any format or medium without formal permission of the copyright holder. The full policy is available online: <http://nrl.northumbria.ac.uk/policies.html>

This document may differ from the final, published version of the research and has been made available online in accordance with publisher policies. To read and/or cite from the published version of the research, please visit the publisher's website (a subscription may be required.)



OPEN

The unaided recovery of marathon-induced serum metabolome alterations

Zinandré Stander¹, Laneke Luies¹, Lodewyk J. Mienie¹, Mari Van Reenen¹, Glyn Howatson^{2,3}, Karen M. Keane², Tom Clifford^{4,5}, Emma J. Stevenson⁴ & Du Toit Loots^{1,6}✉

Endurance athlete performance is greatly dependent on sufficient post-race system recovery, as endurance races have substantial physiological, immunological and metabolic effects on these athletes. To date, the effects of numerous recovery modalities have been investigated, however, very limited literature exists pertaining to metabolic recovery of athletes after endurance races without the utilisation of recovery modalities. As such, this investigation is aimed at identifying the metabolic recovery trend of athletes within 48 h after a marathon. Serum samples of 16 athletes collected 24 h before, immediately after, as well as 24 h and 48 h post-marathon were analysed using an untargeted two-dimensional gas chromatography time-of-flight mass spectrometry metabolomics approach. The metabolic profiles of these comparative time-points indicated a metabolic shift from the overall post-marathon perturbed state back to the pre-marathon metabolic state during the recovery period. Statistical analyses of the data identified 61 significantly altered metabolites including amino acids, fatty acids, tricarboxylic acid cycle, carbohydrates and associated intermediates. These intermediates recovered to pre-marathon related concentrations within 24 h post-marathon, except for xylose which only recovered within 48 h. Furthermore, fluctuations in cholesterol and pyrimidine intermediates indicated the activation of alternative recovery mechanisms. Metabolic recovery of the athletes was attained within 48 h post-marathon, most likely due to reduced need for fuel substrate catabolism. This may result in the activation of glycogenesis, uridine-dependent nucleotide synthesis, protein synthesis, and the inactivation of cellular autophagy. These results may be beneficial in identifying more efficient, targeted recovery approaches to improve athletic performance.

Endurance races (> 5 km) have become increasingly popular over the last decade and primarily include half-marathons (21.1 km), marathons (42.2 km), and ultra-marathons (> 42.2 km)¹. Participation in these events requires extensive preparation including physical and mental conditioning as well as meticulously planned dietary strategies. According to Barnett², optimal athletic performance is only achieved when these principles are accompanied by sufficient system recovery following any endurance activity. “Recovery” comprehensively refers to the process in which an altered biological system, such as the metabolome, reverts to its corresponding pre-perturbed state³. In endurance athletes, metabolic recovery is thought to proceed in a biphasic manner with the initial phase mainly consisting of rapid oxygen, ATP and phosphocreatine replenishment, whereas the second phase entails the slow restoration of innate metabolism adaptations^{2,3}. A popular research approach for identifying and mapping such metabolic adaptations, quantitatively and qualitatively, is collectively referred to as “metabolomics”. Considering that the metabolome provides a direct depiction of the physiological state of an organism at a specific point in time⁴, metabolomics is a formidable tool when investigating the effects of endurance races on the human body, as well as the recovery thereof.

¹Human Metabolomics, Faculty of Natural and Agricultural Sciences, North-West University, Potchefstroom 2531, South Africa. ²Department of Sport, Exercise and Rehabilitation, Faculty of Health and Life Sciences, Northumbria University, Newcastle upon Tyne NE1 8ST, UK. ³Water Research Group, School of Environmental Sciences and Development, North-West University, Potchefstroom 2531, South Africa. ⁴Human Nutrition Research Centre, Institute of Cellular Medicine, Newcastle University, Newcastle upon Tyne, UK. ⁵School of Sport, Exercise and Health Sciences, Loughborough University, Loughborough, UK. ⁶North-West University, Private Bag X6001, Box 269, Potchefstroom Campus Potchefstroom 2520, South Africa. ✉email: dutoit.loots@nwu.ac.za

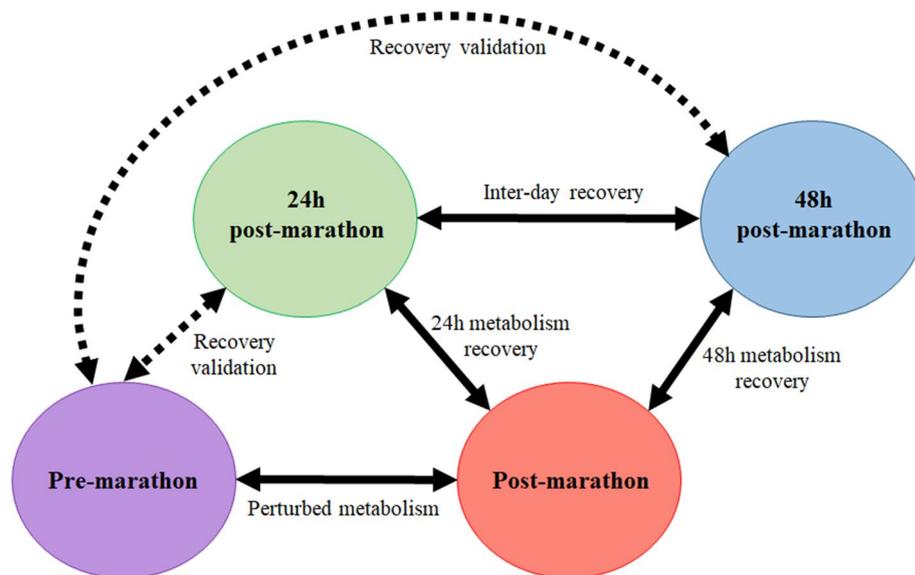


Figure 1. Schematic representation of the sample group comparisons performed during this investigation.

Previous studies investigating the perturbed metabolism of endurance athletes mainly observed elevated concentrations of carbohydrates, fatty acids, and ketones, accompanied by reduced concentrations of amino acids⁵. This is proposed to result from the dramatic metabolic shift between numerous fuel substrate catabolism systems, as well as alternative energy-producing mechanisms (α -oxidation and autophagy) activated during strenuous activity^{6–10}. Additionally, the extensive reduction in fuel substrates was associated with alterations in the urea cycle, purine, steroid and pyrimidine metabolism, reactive oxygen species production pathways^{5,6,9–13}, as well as the gut microbiome^{9,14}.

Given the nature and extent of the physiological, immunological and metabolic effects of endurance exercise on the human body, a substantial amount of research has been directed towards investigating various therapeutic recovery modalities including cryotherapy, massage, heat therapy, stretching, compression, and non-steroidal anti-inflammatory drugs^{2,15–17}, as well as the ingestion of functional foods, such as cherries, pomegranates, and green leafy vegetables¹⁸. However, very limited literature is available regarding the overall recovery of the perturbed metabolism after an endurance race, without the ingestion and/or utilisation of recovery modalities. From these studies, it is evident that during the immediate hours after recovery, myocellular glycogen restoration^{19,20}, as well as phosphocreatine store replenishment and blood lactate removal takes preference^{2,3,21}, while lipids are temporarily further oxidised as an energy source^{19,20}. Considering this, it is suggested that glycogen restoration is attained within 24–48 h after endurance exercise²⁰, whereas amino acids, fatty acids, fatty acid-carnitine conjugates along with tricarboxylic acid (TCA) cycle and pyrimidine/purine pathway intermediates remain perturbed for at least 14 h after endurance running^{5–7,9}.

Considering the current lack of literature regarding the metabolic recovery of marathon athletes over an extended period (48 h), an untargeted two-dimensional gas chromatography time-of-flight mass spectrometry (GCxGC-TOFMS) serum metabolomics approach was used to investigate the unaided (without the ingestion and/or utilisation of recovery modalities) metabolic recovery (to baseline-related concentrations) of 16 marathon athletes within 48 h post-marathon. These results might improve the current knowledge pertaining to the consequence of endurance activity and the time-course of recovery in the days following the event, as well as lead to the identification of new or additional recovery pathways and/or more targeted, perhaps even personalised, recovery approaches.

Results

The GCxGC-TOFMS system detected a total of 1,308 peaks of which 444 peaks were classified as “Analytes/Unknowns”. After removing column and/or reagent-related compounds, as well as following the merging of duplicate peaks of compounds, such as sugars and amino acids, 839 compounds (including the 444 peaks classified as “Analytes/Unknowns” and internal standard) remained. The entire list of 838 compounds (excluding internal standard) was then subjected to the aforementioned processing steps [i.e. quality control coefficient of variation (QC-CV) filter etc.] and further statistical analyses. Considering the aim of this investigation, multiple group comparisons were required (Fig. 1), subsequently leading to numerous statistical outputs, which will be discussed below.

Multivariate statistical output. In order to visualise the overall metabolic shift over the entire time-frame of the study, an analysis of variance (ANOVA) simultaneous component analysis (ASCA) model was constructed with the resulting scores graphically displayed and grouped according to the time component of the current study design (Fig. 2). All scripts used here are readily available in the mentioned toolboxes provided by

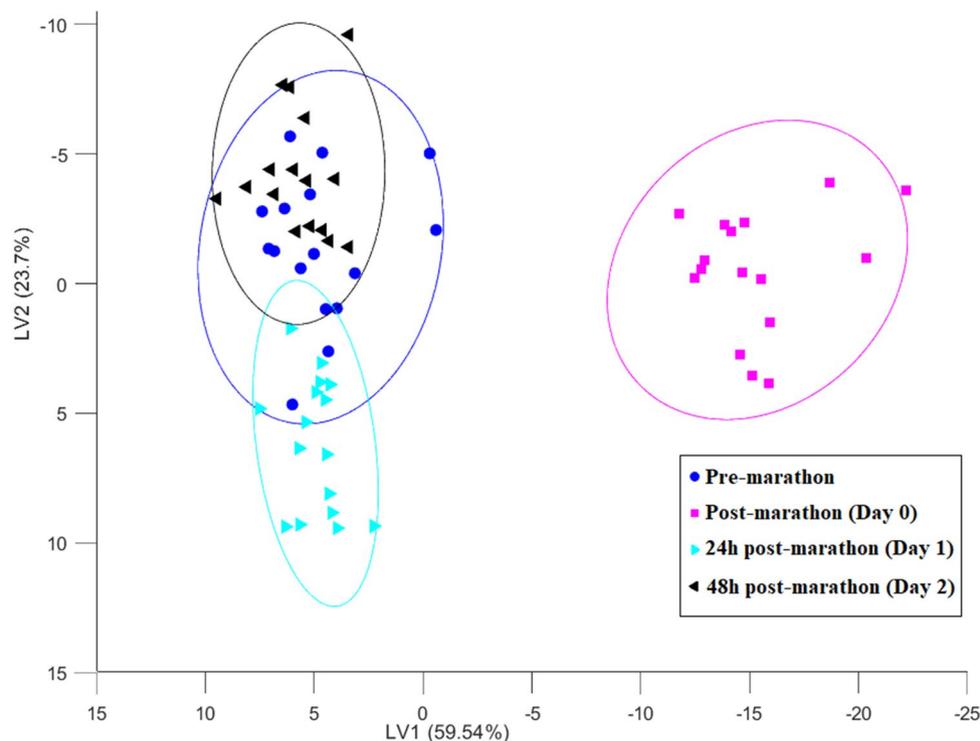


Figure 2. The ASCA plot indicating the natural, time-dependent differentiation of the comparative groups: pre-marathon (denoted by blue/circle), post-marathon (denoted by pink/square), 24 h post-marathon (denoted by turquoise/right-tilted triangle) and 48 h post-marathon (denoted with black/left-tilted triangle). The variance accounted for by each latent variable (LV) is indicated in parenthesis on the respective axes. The ellipsoids represent 95% confidence intervals of each time-points centroid.

Eigenvector Research Inc., with the exception of the ASCA scripts, which can be downloaded from <https://www.bdagroup.nl/content/Downloads/software/software.php>. Based on Fig. 2, it is evident that although a marathon has a dramatic impact on the human serum metabolome these changes recover within 48 h, as indicated by the confined positioning of the 24 h and 48 h post-marathon ellipsoids relative to that of the pre-marathon. Furthermore, in an attempt to inspect specific adaptations between time-points, individual multi-level principal component analysis (ML-PCA) plots for the specified group comparisons (Fig. 1) are presented in Supplementary Fig. S1a–f. The following can be deduced from these figures (S1a–f): (1) As expected, the marathon perturbation (Fig. S1a) induced significant metabolic alterations, which is supported by Stander et al.¹⁰; (2) Clear separation (no overlap) of the recovery time-points (24 h and 48 h post-marathon) and post-marathon (day 0) ellipsoids (Fig. S1b and c) indicate a significant deviation from the post-marathon perturbed metabolic state; (3) The magnitude of metabolome variation presented 24 h and 48 h post-marathon (Fig. S1 f) seems rather minor as these ellipsoids appear relatively grouped (ellipsoid overlap); (4) The confined positioning of the 24 h and 48 h post-marathon ellipsoids relative to the pre-marathon (Fig. S1 d and e) indicate the recovery of the metabolome to a baseline-related state (overlapping ellipsoids). Furthermore, in order to visualise metabolite stratification among the comparative groups, clustering dendrograms were constructed and are included in the supplementary material (Fig. S3). From these figures, it is clear that stratification associations did not correlate with the comparative groups.

However, considering the relatively small sample set and the likelihood of increased false discovery rates (FDR; high dimensionality), along with controlling for model overfitting, multivariate models could not be validated and were therefore not used for variable selection but rather for global metabolic shift visualisation.

Univariate statistical outputs. Considering the aforementioned, variable selection was based on univariate comparisons as a means of determining the specific metabolic adaptations apparent during the course of this investigation while controlling the FDRs and avoiding overfitting. Although repeated measures ANOVA models may have been appropriate, paired t-tests were preferred since the dramatic effect of the marathon would result in a large number of significant metabolites requiring pairwise comparisons as post-hoc tests in any event. Based on this, only metabolites with a Benjamini–Hochberg p value ≤ 0.05 and an effect size d -value ≥ 0.5 (moderate effect) in each group comparison (Fig. 1), were deemed to be statistically significant. This approach yielded a list of 74 metabolite markers, of which 61 were positively annotated (the remaining 13 metabolites were classified as “Unknowns/Analytes”) and are listed in Table 1. A more comprehensive table including average concentrations at each of the comparative time-points has been included in Supplementary Table S1.

Compound name (PubChem ID)	PM/PrM ratio	D1/PM ratio	D2/PM ratio	D2/D1 ratio	D1/PrM ratio	D2/PrM ratio
Carbohydrate and associated intermediates						
Arabitol (439255)	1.36*	0.75*	0.65*	0.87	1.01	0.88
Erythritol (222285)	1.77*	0.61*	0.39*	0.63	1.08	0.68
Glucose (5793)	8.91*	0.17*	0.09*	0.52	1.48	0.77
Mannitol (6251)	1.34	0.78	0.58*	0.75	1.05	0.78
Mannose (18950)	15.72*	0.11*	0.05*	0.49	1.66	0.82
Myo-inositol (440194)	1.10	0.76*	0.63*	0.82	0.84	0.69
Psicose (441036)	1.15	0.45*	0.70	1.56	0.52	0.80
Rhamnose (5460029)	0.62*	1.51	1.86*	1.23	0.94	1.16
Ribose (10975657)	1.16	0.58*	0.89	1.53	0.67	1.03
Sorbose (439192)	1.46	0.69	0.64*	0.94	1.01	0.94
Tagatose (439312)	1.87*	0.56*	0.53*	0.95	1.04	0.99
Talose (441035)	1.69*	0.57*	0.52*	0.91	0.96	0.87
Threonic acid (151152)	1.39*	0.80*	0.78*	0.98	1.11	1.09
Uridine (6029)	1.06	0.61*	1.02	1.66*	0.65	1.08
Xylose (135191)	0.90	0.52*	0.76	1.45	0.47*	0.69
Lipid and associated intermediates						
α -Hydroxyoctanoic acid (94180)	1.49*	0.63*	0.67	1.06	0.94	0.99
α -Linolenic acid (5280934)	1.26*	0.69*	0.74*	1.07	0.87	0.93
β -Hydroxyhexanoic acid (11829482)	3.00*	0.38*	0.31*	0.82	1.13	0.92
5-Dodecenoic acid (5312377)	35.22*	0.02*	0.06*	3.73	0.59	2.19
10-Heptadecenoic acid (5312434)	5.87*	0.12*	0.15*	1.22	0.73	0.89
11,14-Eicosadienoic acid (5282805)	1.94*	0.54*	0.45*	0.83	1.05	0.87
11-Eicosenoic acid (5282768)	3.48*	0.31*	0.37*	1.20	1.06	1.28
Docosahexaenoic acid (445580)	1.17	0.69*	0.80	1.17	0.80	0.94
Eicosapentaenoic acid (446284)	1.30	0.69*	0.74*	1.07	0.90	0.96
Glycerol (753)	5.84*	0.17*	0.19*	1.11	0.99	1.10
Glycerol monopalmitic acid (308463)	1.87*	0.44*	0.55*	1.24	0.82	1.02
Heptadecanoic acid (10465)	1.82*	0.52*	0.52*	1.01	0.94	0.95
Lauric acid (3893)	2.07*	0.57*	0.46*	0.81	1.17	0.95
Methyl oleic acid (9922235)	1.68*	0.48*	0.72	1.51	0.81	1.22
Myristoleic acid (5281119)	12.33*	0.06*	0.08*	1.32	0.70	0.93
Oleic acid (445639)	5.04*	0.21*	0.22*	1.07	1.06	1.13
Palmitic acid (985)	1.67*	0.55*	0.57*	1.04	0.93	0.96
Palmitoleic acid (445638)	6.26*	0.19*	0.22	1.17	1.20	1.40
Pentadecanoic acid (13849)	2.21*	0.51*	0.43*	0.85	1.13	0.96
Ketones and TCA cycle intermediates						
α -Ketoglutaric acid (51)	1.57*	0.58*	0.64*	1.11	0.91	1.01
β -Hydroxybutyric acid (441)	12.98*	0.06*	0.06*	1.12	0.72	0.81
Acetoacetic acid (96)	2.89	0.37*	0.19*	0.51	1.07	0.54
Citric acid (311)	1.46	0.74	0.58*	0.78	1.08	0.85
Fumaric acid (444972)	1.42	0.75	0.68*	0.91	1.06	0.96
Malic acid (525)	1.62*	0.57*	0.52*	0.90	0.93	0.84
Malonic acid (867)	1.37	0.72	0.62*	0.86	0.99	0.85
Amino acid and cholesterol-associated intermediates						
α -Ethylhydracrylic acid (188979)	2.24*	0.48*	0.47*	0.97	1.08	1.05
α -Hydroxybutyric acid (11266)	2.53*	0.52*	0.49*	0.93	1.31	1.23
β -Hydroxyisobutyric acid (87)	15.39*	0.04*	0.06*	1.29	0.67	0.87
β -Hydroxyisovaleric acid (69362)	1.15	0.89	0.75*	0.84	1.02	0.86
5-Pregnen-3 β ,20 α -diol (312224064)	1.93*	0.38*	0.36*	0.96	0.73	0.70
Alanine (5950)	0.64	2.34*	1.86*	0.79	1.51	1.19
Aminomalonic acid (100714)	0.48	2.10*	2.06*	0.98	1.01	0.99
Aspartic acid (5960)	0.60	2.03*	1.36	0.67	1.22	0.82
Glutamic acid (33032)	0.79	1.51*	1.39	0.92	1.19	1.10
Glutaric acid (743)	1.73	1.12	0.52*	0.46	1.94	0.89
Continued						

Compound name (PubChem ID)	PM/PrM ratio	D1/PM ratio	D2/PM ratio	D2/D1 ratio	D1/PrM ratio	D2/PrM ratio
Glycine (750)	0.59	1.79*	1.45	0.81	1.06	0.86
Hydroxyproline (5810)	0.44	2.03	2.68*	1.32	0.89	1.18
Methionine (6137)	0.54	2.17*	1.88*	0.87	1.18	1.02
Phenylalanine (6140)	0.64	1.93*	1.47*	0.76	1.24	0.95
<i>p</i> -Hydroxyphenylacetic acid (127)	2.00	0.51*	0.45*	0.88	1.03	0.90
<i>p</i> -Hydroxyphenyllactic acid (9378)	1.53	0.66*	0.58*	0.88	1.01	0.89
Pyroglutamic acid (7405)	0.71*	1.35*	1.41*	1.04	0.95	1.00
Squalene (638072)	3.39*	0.31*	0.34*	1.07	1.06	1.14
Tyrosine (6057)	0.64	1.78*	1.57*	0.88	1.14	1.01
Valine (6287)	0.53	1.83	2.21*	1.21	0.97	1.18

Table 1. Ratio of concentration fluctuations in the serum metabolite markers (n = 61) best describing the detected variation and metabolic recovery of the athletes within 48 h post-marathon. *PrM* Pre-marathon, *PM* Post-marathon, *D1* 24 h post-marathon, *D2* 48 h post-marathon. Significant *p*- and *d*-values (*).

Discussion

Table 1 depicts the recovery trend of the metabolites not only initially affected by the marathon perturbation, but also the concentration fluctuations of additional metabolites pertinent to the 24 h and 48 h recovery period. In order to ascertain the metabolic recovery mechanisms following a marathon, it is imperative to first consider the immediate effects of a marathon on the serum metabolome of athletes, which was characterised and previously described by Stander et al.¹⁰. Briefly, alterations were mainly observed in metabolic pathways associated with energy production, such as elevated concentrations of various carbohydrates, TCA cycle intermediates, fatty acids and ketone bodies, and reduced concentrations of various amino acids and associated metabolites. The perturbed metabolic state may also be linked to imbalanced redox potential, the saturation of traditional fatty acid catabolic pathways, possible insulin secretion hampering, and the activation of cellular autophagy to provide additional fuel substrates.

A detailed description of the recovery (to pre-marathon related concentrations) of the identified metabolites are discussed in a grouped-manner, based on metabolic alterations of the recovery time-points (24 h and 48 h post-marathon; day 1 and day 2) relative to the perturbed metabolic state (day 0; immediately post-marathon) and is schematically summarised in Fig. 3.

Carbohydrate metabolism. Carbohydrates, glucose in particular, are considered to be the primary fuel substrates used by the body during physical activity. Stander et al.¹⁰ and Lewis et al.⁶ reported severely elevated serum concentrations of various carbohydrates immediately after a marathon (day 0). In this metabolomics investigation, it is clear that the majority of the serum carbohydrates (glucose and mannose) and associated metabolites (uridine, ribose, glycerol, and myo-inositol) returned to baseline (pre-marathon) concentrations within 24 h post-marathon (Table 1), with the exception of xylose (also dietary-related). This can be ascribed to the restoration of the glycogen stores, which were perceived as "depleted" in the initial stages of the marathon¹⁰.

Moreover, substantial reductions in uridine concentrations were observed within 24 h post-marathon (Table 1). Uridine is a nucleoside formed via the conjugation product of ribose and uracil and is a key component in RNA and subsequently protein synthesis²². This, in conjunction with the reduced ribose, indicated possible increased nucleotide production via the activation of the pyrimidine salvation metabolism. The pyrimidine salvation metabolism refers to the adaptive mechanism by which nucleotides can be synthesised using pre-formed nucleosides and nucleobases, usually following a nucleotide degradation period^{5,22}, such as the extensive protein catabolism induced by a marathon^{6,7,10}. These recycled nucleotides can also interact with lipids to form pyrimidine nucleotide–lipid conjugates, a key component in cell membrane structures²². Considering this, the slight reduction in uridine concentrations observed 24 h after the completion of the marathon, may also be associated with the restoration of damaged cellular membranes, which occur as a result of strenuous exercise¹⁰. This is further substantiated by the reduction/recovery of myo-inositol concentrations within 24 h post-marathon, as it is a well-known constituent of glycerophospholipids²³.

Additionally, the *de novo* pyrimidine synthesis pathway is dependent on various elements of the pentose phosphate pathway²², particularly ribose-5-phosphate and phosphoribosyl pyrophosphate²⁴. Since these are interlinking substrates for the production of erythrose-4-phosphate²⁴, the recovery of erythritol (produced from the latter)²⁵ not only suggests the activation of glycogenesis via the utilisation of pentose phosphate pathway intermediates, but further supports the activation of the abovementioned nucleotide repair mechanisms. Furthermore, since the majority of the erythritol is absorbed rather than metabolized²⁶, as well as the fact that it is a well-known alternative sugar added to low-calorie foods²⁷ frequently ingested by athletes, the alterations of this metabolite 24 h and 48 h post-marathon may also be due to reduced ingestion of such foods by the athletes during the recovery period. Similarly, several of the metabolites in Table 1, including mannitol, psicose, sorbose^{27,28}, tagatose²⁹, and xylose²⁴, are natural constituents of fruit and vegetables, Food and Drug Administration approved alternative bulk sweeteners, and/or the products of metabolised sweeteners such as threonic acid³⁰, and arabinol via arabinose³¹. Considering that endurance races reportedly alter gut microbiome composition, as well as increase intestinal permeability resulting in gut microbiome (GM) metabolite migration³², reductions in several

24 h after the marathon (Table 1), support the aforementioned glycaemic flux recovery. This concurs with Nieman et al.⁷, who reported that these pathways remain perturbed for approximately 14 h post-exercise. Additionally, reduced concentrations of ketones, particularly β -hydroxybutyric acid, support the restoration of the previously mentioned redox imbalance^{10,38} and suggest an enhanced metabolic clearance rate⁴¹ in athletes within 24 h post-marathon. β -Hydroxybutyric acid is also a signalling molecule serving as a regulator of gene expression and adaptive responses via the inhibition of class I histone deacetylases, rendering this metabolite invaluable to the recovery process⁴¹.

Amino acid metabolism. In accordance with the carbohydrates and lipids profile, various amino acids (alanine, aspartic acid, glycine, glutamic acid, methionine, phenylalanine, and tyrosine) and associated metabolites (glutaric acid, *p*-hydroxyphenylacetic acid, *p*-hydroxyphenyllactic acid, pyroglutamic acid, and hydroxyproline) recovered to baseline concentrations within 24 h after the completion of the marathon (Table 1), indicating a reduced need for amino acid catabolism during this recovery phase. However, the slight reductions in these amino acids that occur between 24 and 48 h of recovery could be an indication of protein synthesis, as substantiated by the aforementioned uridine/ribose profile (Table 1). Contrarily, slight elevations in recovery time-point hydroxyproline may be indicative of delayed post-marathon collagen catabolism, thus supporting previous literature⁴². Furthermore, the reduced concentrations of α -hydroxybutyric acid (a threonine/methionine metabolism intermediate) within 24 h post-marathon compared to the elevated post-marathon (Day 0) profile correlates with reduced amino acid catabolism and further suggests recovery of previously reported transiently hampered insulin secretion¹⁰, glucose absorption, and imbalance redox potential⁴³.

In addition to the proteinogenic amino acids, elevated concentrations of the branched-chain amino acid (BCAA) valine, in conjunction with reduced BCAA-associated catabolism intermediates (Table 1) such as β -hydroxyisobutyric acid (valine metabolism), α -ethylhydracrylic acid (allo-isoleucine metabolism), and β -hydroxyisovaleric acid (leucine metabolism), are indicative of a reduction in BCAA catabolism. BCAAs are reportedly one of the major factors regulating mammalian target of rapamycin complex 1 (mTOR1), a multi-protein complex responsible for the activation of numerous biological processes imperative to metabolic recovery, including cell proliferation, gene expression for lipid and protein synthesis, as well as the inhibition of cellular autophagy⁴⁴. The proposed activation of these recovery mechanisms is further substantiated by the aforementioned uridine/ribose profile (Table 1), as well as the reduction in various cell membrane-associated intermediates, such as fatty acids (particularly OCFAs), myo-inositol and cholesterol intermediates (squalene and 5-pregnen-3 β ,20 α -diol) observed within 24 h post-marathon. Reduced concentrations of cholesterol intermediates within 24 h after the marathon in comparison to the post-marathon (day 0) concentrations also suggest inactivation of steroidogenesis⁴⁵, which was previously proposed to be induced as a result of endurance activity^{10,37}.

Aminomalonic acid was also observed to recover to pre-marathon concentrations within 24 h post-marathon. Although this metabolite has been associated with oxidative damage to protein/amino acid residues⁴⁶, it has also been isolated from *Escherichia coli*⁴⁷. As such, its precise mechanism of action pertaining to the metabolic recovery process remains to be elucidated.

Conclusion

The results of this investigation suggest that the serum metabolome of athletes largely recovers to baseline (pre-marathon) status within 24 h, with total recovery evident within 48 h post-marathon without the intervention of recovery aids. This can be ascribed to a reduction in ATP requirements during passive recovery, which in turn allows for the activation of anabolic recovery mechanisms or pathways. The possible cascade of events resulting in total metabolome recovery within 24/48 h post-marathon are: (a) reduced energy requirements promoting the restoration of the imbalanced redox potential and intracellular glycaemic flux, thus stimulating glycogenesis. (b) In turn, this further reduces the need for alternative fuel substrate catabolism (of lipids and amino acids) and ketogenesis, (c) activating additional cellular repair mechanisms such as the uridine-dependent nucleotide salvage pathway, mTOR1 stimulation (subsequently inactivating cellular autophagy), as well as possible protein and lipid synthesis/fatty acid re-esterification. Moreover, GM-associated markers emphasize the close metabolic interaction between man and microbe.

Limitations and future prospects

Possible limitations of this investigation may include the inevitable confounder of human genotype/phenotype (sex, age and ethnicity) variation, as well as individual athlete dietary variation. Although the athletes' dietary intake was recorded throughout the investigation, adding dietary restrictions to the already refrained use of recovery aids would be considered too much of an interference to the personalised regimens of these athletes. However, in an attempt to circumvent the variability in participant demographics (age, sex, fitness, running experience, etc.), the study design included repeated measures of the same individuals and the use of statistical methods that consider the dependence between observations (therefore each participant becomes their own control)^{48,49}. Nevertheless, such individual variation allows for a higher level of robustness in the results, which provides a better representation of the true nature of the activated recovery mechanisms as the body attempts to reach homeostasis.

Although the current investigation provided a holistic (untargeted) representation of the metabolome recovery after a marathon, future studies may consider applying more targeted approaches based on these results to possibly identify more personalised/efficient recovery modalities. Additionally, since this investigation was performed on a relatively small cohort, future investigations consisting of larger cohorts and/or serum samples obtained at additional time-points (e.g. 12 h, 36 h, and 72 h post-marathon) are required to validate these findings, ultimately allowing for better metabolic mapping of the in-between stages via the identification of additional

Participant demographical information	Average \pm standard deviation
Age (years)	39 \pm 12
Pre-marathon athlete weight (kg)	72.2 \pm 11.9
Post-marathon athlete weight (kg)	70.8 \pm 11.7
Running experience (years)	8 \pm 7
Marathon experience (races)	21 \pm 40
Finishing time (hh:mm:ss)	04:30:25 \pm 00:36:48

Table 2. Participant demographical information and endurance race experience history.

metabolite markers pertinent to the metabolic recovery trend of endurance athletes. Considering that metabolic differences correlating with gender heterogeneity could not be positively annotated due to limited cohort size, future studies may consider including larger gender-specific cohorts. Lastly, since GM metabolism seems to be an integral part of athletic performance, more targeted analysis of the associated metabolites or using an alternative biological medium (such as faecal matter, urine, etc.) could provide a more perspicuous understanding of the importance of the GM in terms of metabolome recovery.

Methods and materials

Participants and the marathon race. The participants ($n = 16$; 10 males and 6 females) were selected at random from the pool of athletes competing in the Druridge Bay Marathon in Newcastle (UK). The eligibility of the selected participants was assessed by means of a health and dietary questionnaire, after which those individuals with food allergies, cardiovascular complications, musculoskeletal disorders/injuries, or those receiving anti-inflammatory treatment were excluded from the study. In addition, female participants completed a menstruation cycle questionnaire. Considering the ethical implications concerning ethnicity classifications of participants when it is not central to the research topic, as well as the fact that this study was not biased towards any ethnic groups, nor were any of the aims a comparison of ethnicity with respect to the marathon induced changes to the human metabolome, athlete ethnicity was not recorded. All athletes were instructed to abstain from using any recovery modalities/therapies such as cryotherapy, pressure garments, foam rolling, active recovery (moderate exercise bouts), NSAID, vitamins, etc. during the recovery period of this investigation. A summary of participant characteristics is presented in Table 2, while a summary of age, sex, and overall performance distribution of the athletes are presented in Table S2 (Supplementary information). The Druridge marathon entails running four laps around the Country Park (Northumberland coast, UK), which mainly consists of grassy, concrete mixed terrain. Approximately 1.6 km of each lap consists of soft sand (± 6.4 km)⁵⁰ and may contribute to slower finishing times (Table 2).

Sample collection. This investigation forms part of a larger, double-blinded placebo study (Supplementary Fig. S2), in which 31 athletes participating in the 2016 Druridge Bay Marathon either received multiple beetroot ($n = 15$ athletes) or placebo ($n = 16$ athletes) supplements, over a period of 48 h. Blood samples ($n = 124$ samples) were collected 24 h before (as a fasting, yet hydrated, baseline/pre-marathon), immediately after (within 30 min; post-marathon [day 0]), as well as 24 h (day 1 post-marathon) and 48 h (day 2 post-marathon) post-marathon by means of antecubital (basilica vein) venepuncture^{10,50}. Venous sampling was limited to only four sampling intervals considering the impact and inconvenience of increased blood sampling, which would have led to increased participant drop-out. These samples were procured from the Faculty of Health and Life Sciences, Department of Sport, Exercise and Rehabilitation at Northumbria University in Newcastle upon Tyne (UK), and transported to the Laboratory of Infectious and Acquired Diseases, Focus Area: Human Metabolomics at the North-West University in Potchefstroom, as previously described by Stander et al.¹⁰. Considering the aim of the current investigation, only the serum samples ($n = 64$) of the 16 athletes that ingested placebo supplements were included in this study. Furthermore, since this study only focuses on the metabolic adaptations presented during a 48 h recovery period, information regarding additional clinical serum measurements are not included, but are available in a complementary physiological publication based on the same cohort⁵⁰.

Sample extraction and derivatisation. As previously described in detail by Stander et al.¹⁰, all samples along with pooled quality control (QC) samples (pooled 50 μ L of each sample) were subjected to a total metabolome extraction. As part of this process, smaller aliquots of the samples (50 μ L) were prepared, before adding 50 μ L of the internal standard (3-phenylbutyric acid; 0.45 μ g/mL) dissolved in a chloroform:methanol:water (1:3:1) solution. The latter serves as a metabolite conservation step to ensure that larger polar and non-polar compounds and associated metabolites (e.g. amino acids, fatty acids, purines, pyrimidines, alkenes, alkanes etc.) can be analysed using the GCxGC-TOFMS. Protein precipitation of the samples commenced by adding 300 μ L of ice-cold acetonitrile, where after each vial was subjected to a 2 min mixing step (REAX D-91126 vortex; Heidolph Instruments GmbH & Co.KG, Schwabach, Germany) and 10 min centrifugation at 4,000 rpm. The supernatants were extracted, transferred to a clean GC-MS vial, and dried with a light stream of nitrogen gas for 45 min, whilst being housed in a heating block set at 40 $^{\circ}$ C. Once the solvent evaporated, 25 μ L of methoxamine hydrochloride dissolved in pyridine (15 mg/mL) was added to each sample and incubated for 90 min at

50 °C. Lastly, 40 µL of BSTFA enriched with 1% TCMS was added, and samples were left to derivatise/silylate for 60 min at 60 °C. Derivatized samples were then transferred to a clean glass GC–MS vial containing a vial insert.

GCxGC-TOFMS analysis. Samples were injected into the Pegasus 4D GCxGC-TOFMS analytical apparatus (LECO Africa (Pty) Ltd, Johannesburg, South Africa) in a randomised order by using a Gerstel 104 autosampler tray (GmbH and co. KG, Mülheim van der Ruhr, Germany), with QC samples preceding approximately every 7th sample for quality assurance purposes. Each sample (1 µL) was injected using a 1:3 split ratio with a constant injector temperature of 270 °C and a carrier gas (purified helium) flow rate of 1 mL/min. The primary oven containing a Restek Rxi-5MS capillary column (30 m; 0.25 µm diameter and 0.25 µm film thickness) was set to operate at an initial temperature of 70 °C before incrementally (4 °C/min) increasing to a final temperature of 300 °C (maintained for 2 min). In addition, the secondary oven, fitted with a Restek Rxi-17 capillary column (1 m; 0.25 µm diameter and 0.25 µm film thickness), was programmed with an initial temperature of 85 °C which was set to incrementally increase with 4.5 °C/min until a final temperature of 300 °C was reached. The thermal modulator, amid the two columns, was set to intermittently pulse streams of hot and cold nitrogen gas for approximately 0.5 s (every 3 s). Mass spectra (ms) of ions (50–800 *m/z* range) was acquired at a rate of 200 ms/s, while the ms of the first 400 s of the run was discarded to exclude extraction solvent elution. However, the latter was not removed from the time axis of the primary column to ensure compound retention time validity. The transfer line and ion source were set to function at 270 °C and 220 °C, respectively, with a detector voltage of 1,600 V and ion filament bias of –70e V. After analytical analysis, the data generated was processed using the ChromaTOF software version 4.7.1 (LECO Corporation), which included ms deconvolution, peak alignment and peak annotation via mass fragment pattern and retention time comparison with in-house libraries compiled from previously injected standards⁵¹. ChromaTof parameter settings included peak widths of 9 and 0.2 for the 1st and 2nd dimension respectively, and a signal-to-noise ratio of 100, with a maximum of 100,000 unknown peaks allowed (to remove bias). All metabolites with molecular masses between 50 and 800 Da were detected with a minimum of 70% similarity match to the in-house libraries for identification. The experimental procedure performed was in accordance with proposed guidelines and regulations, as approved by the North-West University (South Africa) and forms part of the laboratory's standard operating procedures (SOP number: HM-MET-056).

Statistical analysis and group comparison. Prior to data analysis, pooled QC samples were compared to ensure data validity before interpretation. Data analyses commenced with compound normalisation in relation to an internal standard and all identified analytical contaminants/plasticizers/column related metabolites were removed. The data was then subjected to numerous data clean-up steps, including a QC-CV filter (retaining metabolites with CV ≤ 30), QC drift/batch correction (using a pooled QC sample), zero-value replacement (with random, small values along the tail distribution of the data), 50% zero value filter, log transformation and auto-scaling^{48,52–54}. Hereafter the data was subjected to a multi-statistical approach consisting of two main objectives i.e. (1) obtain a global pattern of the metabolic profile of athletes over time and to some extent validate the design against the logical result, as well as (2) determine specific metabolic adaptations between time-points (Fig. 1) to distinguish between sources of variation due to early (day 1 vs day 0) and late (day 2 vs day 0) stages recovery; from full recovery (day 2 vs pre-marathon). As such, statistical analyses consisted of both paired univariate and multivariate approaches, using MATLAB⁵⁵ with a PLS toolbox⁵⁶. In terms of the multivariate analyses, an ML-PCA were used to inspect specific time-point differences in the multivariate space, whereas an ASCA was employed to summarise the overall metabolic perturbation observed over the entire timeframe of the study⁵⁷. Both these methods take the dependence between observations into account. Furthermore, univariate data analyses included a paired *t*-test, corrected for multiple testing by controlling the FDR (limiting to 5%) using the Benjamini–Hochberg procedure⁵⁸, and a dependent effect size test to respectively assess the statistical and practical significance of each metabolite⁴⁸.

Considering the inevitable confounder of inter-individual variation, a repeated measures statistical design was employed to control for the latter as each athlete's metabolic variation is evaluated relative to him/herself. That said, statistical approaches for paired data may not correct for all confounders and as such this design aimed to ensure equal representation across different confounders as indicated in Table S2 (Supplementary information). Although more complex models were considered for variable selection (i.e. repeated measures ANOVA), these models would not be applicable due to the small sample size and were therefore excluded.

Ethical approval. Ethical approval for this investigation, conducted according to the Declaration of Helsinki and International Conference on Harmonization Guidelines, was obtained from the Research Ethics Committee of the Faculty of Health and Life Sciences at the Northumbria University in Newcastle upon Tyne, UK (reference number: HLSTC120716). Informed, written consent was obtained from all individuals included in the study.

Data availability

As the current investigation is part of a larger collaboration study consisting of multiple aims, the dataset of this investigation is not yet publicly available but can be acquired from the corresponding author on reasonable request. The authors declare that all the results included in this study have been presented clearly, honestly and without fabrication, falsification, or inappropriate data manipulation.

Received: 29 August 2019; Accepted: 2 June 2020

Published online: 06 July 2020

References

- Lippi, G. *et al.* Troponin I measured with a high sensitivity immunoassay is significantly increased after a half marathon run. *Scand. J. Clin. Lab. Invest.* **72**, 467–470. <https://doi.org/10.3109/00365513.2012.697575> (2012).
- Barnett, A. Using recovery modalities between training sessions in elite athletes. *Sports Med.* **36**, 781–796 (2006).
- Tomlin, D. L. & Wenger, H. A. The relationship between aerobic fitness and recovery from high intensity intermittent exercise. *Sports Med.* **31**, 1–11. <https://doi.org/10.2165/00007256-200131010-00001> (2001).
- Heaney, L. M., Deighton, K. & Suzuki, T. Non-targeted metabolomics in sport and exercise science. *J. Sports Sci.* <https://doi.org/10.1080/02640414.2017.1305122> (2017).
- Davison, G. *et al.* Metabolomic response to acute hypoxic exercise and recovery in adult males. *Front. Physiol.* **9**, 1682. <https://doi.org/10.3389/fphys.2018.01682> (2018).
- Lewis, G. D. *et al.* Metabolic signatures of exercise in human plasma. *Sci. Transl. Med.* **2**, 33–37. <https://doi.org/10.1126/scitranslmed.3001006> (2010).
- Nieman, D. C., Shanelly, R. A., Gillitt, N. D., Pappan, K. L. & Lila, M. A. Serum metabolic signatures induced by a three-day intensified exercise period persist after 14 h of recovery in runners. *J. Proteome Res.* **12**, 4577–4584. <https://doi.org/10.1021/pr400717j> (2013).
- Waśkiewicz, Z. *et al.* Acute metabolic responses to a 24-h ultra-marathon race in male amateur runners. *Eur. J. Appl. Physiol.* **112**, 1679–1688. <https://doi.org/10.1007/s00421-011-2135-5> (2012).
- Daskalaki, E., Easton, C. & Watson, D. G. The application of metabolomic profiling to the effects of physical activity. *Curr. Metab.* **2**, 233–263. <https://doi.org/10.2174/2213235x03666150211000831> (2014).
- Stander, Z. *et al.* The altered human serum metabolome induced by a marathon. *Metabolomics* <https://doi.org/10.1007/s11306-018-1447-4> (2018).
- Turer, A. T. *et al.* Increases in myocardial workload induced by rapid atrial pacing trigger alterations in global metabolism. *PLoS ONE* **9**, 1–9. <https://doi.org/10.1371/journal.pone.0099058> (2014).
- Kujala, U. M. *et al.* Long-term leisure-time physical activity and serum metabolome. *Circulation* **127**, 340–348. <https://doi.org/10.1161/CIRCULATIONAHA.112.105551> (2013).
- Clifford, T. *et al.* A pilot study investigating reactive oxygen species production in capillary blood after a marathon and the influence of an antioxidant-rich beetroot juice. *Appl. Physiol. Nutr. Metab.* **43**(3), 303–306 (2017).
- Mach, N. & Fuster-Botella, D. Endurance exercise and gut microbiota: a review. *J. Sport Health Sci.* **6**, 179–197. <https://doi.org/10.1016/j.jshs.2016.05.001> (2017).
- Howatson, G. & Van Someren, K. A. The prevention and treatment of exercise-induced muscle damage. *Sports Med.* **38**, 483–503 (2008).
- Wilson, L. J. *et al.* Recovery following a marathon: a comparison of cold water immersion, whole body cryotherapy and a placebo control. *Eur. J. Appl. Physiol.* **118**, 153–163. <https://doi.org/10.1007/s00421-017-3757-z> (2018).
- Brown, F. *et al.* Compression garments and recovery from exercise: a meta-analysis. *Sports Med.* **47**, 2245–2267. <https://doi.org/10.1007/s40279-017-0728-9> (2017).
- Harty, P. S., Cottet, M. L., Malloy, J. K. & Kerksick, C. M. Nutritional and supplementation strategies to prevent and attenuate exercise-induced muscle damage: a brief review. *Sports Med. Open* **5**, 1–17. <https://doi.org/10.1186/s40798-018-0176-6> (2019).
- Bielinski, R., Schutz, Y. & Jequier, E. Energy metabolism during the postexercise recovery in man. *Am. J. Clin. Nutr.* **42**, 69–82. <https://doi.org/10.1093/ajcn/42.1.69> (1985).
- Egan, B. & Zierath, J. R. Exercise metabolism and the molecular regulation of skeletal muscle adaptation. *Cell Metab.* **17**, 162–184. <https://doi.org/10.1016/j.cmet.2012.12.012> (2013).
- Bishop, P. A., Jones, E. & Woods, A. K. Recovery from training: a brief review. *J. Strength Cond. Res.* **22**, 1015–1024. <https://doi.org/10.1519/JSC.0b013e31816eb518> (2008).
- Yamamoto, T. *et al.* Biochemistry of uridine in plasma. *Clin. Chim. Acta* **412**, 1712–1724. <https://doi.org/10.1016/j.cca.2011.06.006> (2011).
- Christie, W. W. & Han, X. *Lipid Analysis: Isolation, Separation, Identification and Lipidomic Analysis* 4th edn, Vol. 24 (Woodhead Publishing Ltd, Cambridge, 2012).
- Salway, J. G. *Medical Biochemistry at a Glance* 3rd edn, 26–127 (Wiley, Berlin, 2012).
- Hootman, K. C. *et al.* Erythritol is a pentose-phosphate pathway metabolite and associated with adiposity gain in young adults. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 4233–4240. <https://doi.org/10.1073/pnas.1620079114> (2017).
- Boesten, D. M. *et al.* Multi-targeted mechanisms underlying the endothelial protective effects of the diabetic-safe sweetener erythritol. *PLoS ONE* **8**, e65741. <https://doi.org/10.1371/journal.pone.0065741> (2013).
- Chattopadhyay, S., Raychaudhuri, U. & Chakraborty, R. Artificial sweeteners—a review. *J. Food Sci. Technol.* **51**, 611–621. <https://doi.org/10.1007/s13197-011-0571-1> (2014).
- Yamada, T. *et al.* Dietary D-sorbitol decreases serum insulin levels in growing Sprague–Dawley rats. *J. Nutr. Sci. Vitaminol.* **60**, 297–299. <https://doi.org/10.3177/jnsv.60.297> (2014).
- Guerrero-Wyss, M., Duran Aguero, S. & Angarita Davila, L. D-tagatose is a promising sweetener to control glycaemia: a new functional food. *Biomed. Res. Int.* <https://doi.org/10.1155/2018/8718053> (2018).
- Thomas, M. & Hughes, R. E. A relationship between ascorbic acid and threonic acid in guinea-pigs. *Food Chem. Toxicol.* **21**, 449–452. [https://doi.org/10.1016/0278-6915\(83\)90101-1](https://doi.org/10.1016/0278-6915(83)90101-1) (1983).
- Kumdam, H., Murthy, S. N. & Gummadi, S. N. Arabitol production by microbial fermentation—biosynthesis and future applications. *Int. J. Sci. Appl. Res.* **1**, 1–12 (2014).
- Rankin, A., O'Donovan, C., Madigan, S. M., O'Sullivan, O. & Cotter, P. D. 'Microbes in sport'—the potential role of the gut microbiota in athlete health and performance. *Br. J. Sports Med.* **51**, 698–699. <https://doi.org/10.1136/bjsports-2016-097227> (2017).
- Terami, Y. *et al.* Production of L-allose and D-talose from L-psicose and D-tagatose by L-ribose isomerase. *Biosci. Biotechnol. Biochem.* **79**, 1725–1729. <https://doi.org/10.1080/09168451.2015.1038215> (2015).
- MacLaren, D. & Morton, J. *Biochemistry for Sport and Exercise Metabolism* 1st edn, 145–225 (Wiley, New York, 2012).
- Baker, J. S., McCormick, M. C. & Robergs, R. A. Interaction among skeletal muscle metabolic energy systems during intense exercise. *J. Nutr. Metab.* <https://doi.org/10.1155/2010/905612> (2010).
- Kimber, N. E., Heigenhauser, G. J., Spriet, L. L. & Dyck, D. J. Skeletal muscle fat and carbohydrate metabolism during recovery from glycogen-depleting exercise in humans. *J. Physiol.* **548**, 919–927. <https://doi.org/10.1113/jphysiol.2002.031179> (2003).
- Daskalaki, E. *et al.* A study of the effects of exercise on the urinary metabolome using normalisation to individual metabolic output. *Metabolites* **5**, 119–139. <https://doi.org/10.3390/metabo5010119> (2015).
- Vidal, K., Robinson, N. & Ives, S. J. Exercise performance and physiological responses: the potential role of redox imbalance. *Physiol. Rep.* **5**, 1–15. <https://doi.org/10.14814/phy2.13225> (2017).
- Singh, R. & Cuervo, A. M. Autophagy in the cellular energetic balance. *Cell Metab.* **13**, 495–504. <https://doi.org/10.1016/j.cmet.2011.04.004> (2011).
- Jenkins, B. J. *et al.* Odd chain fatty acids; new insights of the relationship between the gut microbiota, dietary intake, biosynthesis and glucose intolerance. *Sci. Rep.* **7**, 1–8. <https://doi.org/10.1038/srep44845> (2017).

41. Evans, M., Cogan, K. E. & Egan, B. Metabolism of ketone bodies during exercise and training: physiological basis for exogenous supplementation. *J. Physiol.* **595**, 2857–2871. <https://doi.org/10.1113/JP273185> (2017).
42. Tofas, T. *et al.* Plyometric exercise increases serum indices of muscle damage and collagen breakdown. *J. Strength Cond. Res.* **22**, 490–496 (2008).
43. Gall, W. E. *et al.* Alpha-hydroxybutyrate is an early biomarker of insulin resistance and glucose intolerance in a nondiabetic population. *PLoS ONE* **5**, e10883. <https://doi.org/10.1371/journal.pone.0010883> (2010).
44. Laplante, M. & Sabatini, D. M. mTOR signaling at a glance. *J. Cell Sci.* **122**, 3589–3594. <https://doi.org/10.1242/jcs.051011> (2009).
45. Ebner, M. J., Corol, D. I., Havlikova, H., Honour, J. W. & Fry, J. P. Identification of neuroactive steroids and their precursors and metabolites in adult male rat brain. *Endocrinology* **147**, 179–190. <https://doi.org/10.1210/en.2005-1065> (2006).
46. Bujak, R. *et al.* New biochemical insights into the mechanisms of pulmonary arterial hypertension in humans. *PLoS ONE* **11**, 1–14. <https://doi.org/10.1371/journal.pone.0160505> (2016).
47. Van Buskirk, J. J., Kirsch, W. M., Kleyer, D. L., Barkley, R. M. & Koch, T. H. Aminomalonic acid: identification in *Escherichia coli* and atherosclerotic plaque. *Proc. Natl. Acad. Sci. USA* **81**, 722–725 (1984).
48. Luies, A. *Discovering Statistics Using SPSS* 3rd edn. (Sage, Beverley Hills, 2011).
49. Howitt, D. & Cramer, D. *Introduction to Research Methods in Psychology* 3rd edn. (Pearson Education Limited, London, 2011).
50. Clifford, T. *et al.* Minimal muscle damage after a marathon and no influence of beetroot juice on inflammation and recovery. *Appl. Physiol. Nutr. Metab.* **42**, 263–270. <https://doi.org/10.1139/apnm-2016-0525> (2017).
51. Luies, L. & Loots, D. T. Tuberculosis metabolomics reveals adaptations of man and microbe in order to outcompete and survive. *Metabolomics* **12**, 1–9. <https://doi.org/10.1007/s11306-016-0969-x> (2016).
52. Wang, S.-Y., Kuo, C.-H. & Tseng, Y. J. Batch normalizer: a fast total abundance regression calibration method to simultaneously adjust batch and injection order effects in liquid chromatography/time-of-flight mass spectrometry-based metabolomics data and comparison with current calibration methods. *Anal. Chem.* **85**, 1037–1046. <https://doi.org/10.1021/ac302877x> (2013).
53. Fernandez, J. A. M., Vida, C. B. & Glahn, V. P. In *Data Analysis, Classification, and Related Methods* (eds Kiers, H. A. L., Rasson, J.-P., Groenen, P. J. F. & Schader, M.) (Springer, Berlin 2000), pp. 155–160.
54. Chaleckis, R., Murakami, I., Takada, J., Kondoh, H. & Yanagida, M. Individual variability in human blood metabolites identifies age-related differences. *Proc. Natl. Acad. Sci. USA* **113**, 4252–4259 (2016).
55. MATLAB and Statistics toolbox v. 2012b (The MathWorks Inc., Natick, 2012).
56. PLS_Toolbox 8.2.1 v. 8.2.1 (Eigenvector Research Inc., Manson, 2016).
57. Smilde, A. K. *et al.* ANOVA-simultaneous component analysis (ASCA): a new tool for analyzing designed metabolomics data. *Bioinformatics* **21**, 3043–3048. <https://doi.org/10.1093/bioinformatics/bti476> (2005).
58. Yoav, M. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc.* **57**, 289–300 (1995).

Acknowledgements

The authors would like to sincerely thank Ms Derylyze Beukes for sample extraction/analysis, and Prof. Nico L. Smit for initiating/managing the institutional collaboration. This work is based on the research supported in part by the National Research Foundation of South Africa (Grant Nos. 116703, 120358). The opinions expressed, and conclusions derived are those of the authors and are not necessarily those of the NRF. The authors have no other specific funding to report. No writing assistance was utilised in the production of this manuscript.

Author contributions

The concept and study design were formulated by D.T.L., Z.S., G.H., T.C., K.M.K., and E.J.S. The samples were acquired from the Northumbria University (G.H., T.C., K.M.K., and E.M.S.) as part of a collaboration agreement with the North-West University. Statistical analyses were performed by M.R. while manuscript drafting, data analysis, and interpretation were performed by Z.S. D.T.L., L.L., and L.J.M. contributed to in-depth data interpretation and the manuscript was repeatedly reviewed by D.T.L. and L.L., of which L.L. provided substantial insights into structural (format) editing. All of the authors reviewed and approved the final draft of this manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41598-020-67884-9>.

Correspondence and requests for materials should be addressed to D.T.L.

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

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



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

© The Author(s) 2020