Lipidomic profiles, lipid trajectories and clinical biomarkers in female elite endurance athletes

Tibor V. Varga1,2, Ashfaq Ali3,4, Jose A. R. Herrera1, Linda L. Ahonen3,4, Ismo M. Mattila3, Naba H. Al-Sari3, Cristina Legido-Quigley3,6, Sven Skouby7,8, Søren Brunak1,10* & Åsa B. Tornberg9,10*

We assessed whether blood lipid metabolites and their changes associate with various cardiometabolic, endocrine, bone- and energy-related comorbidities of Relative Energy Deficiency in Sport (RED-S) in female elite endurance athletes. Thirty-eight Scandinavian female elite athletes underwent a day-long exercise test. Five blood samples were obtained during the day - at fasting state and before and after two standardized exercise tests. Clinical biomarkers were assessed at fasting state, while untargeted lipidomics was undertaken using all blood samples. Linear and logistic regression was used to assess associations between lipidomic features and clinical biomarkers. Overrepresentations of findings with $P < 0.05$ from these association tests were assessed using Fisher’s exact tests. Self-organizing maps and a trajectory clustering algorithm were utilized to identify informative clusters in the population. Twenty associations $PFDR < 0.05$ were detected between lipidomic features and clinical biomarkers. Notably, cortisol demonstrated an overrepresentation of associations with $P < 0.05$ compared to other traits ($PFisher = 1.9 \times 10^{-14}$). Mean lipid trajectories were created for 201 named features for the cohort and subsequently by stratifying participants by their energy availability and menstrual dysfunction status. This exploratory analysis of lipid trajectories indicates that participants with menstrual dysfunction might have decreased adaptive response to exercise interventions.

Relative Energy Deficiency in Sport (RED-S) represents a metabolically altered state that is associated with impaired metabolic rate and bone health, decreased immunity and protein synthesis, and cardiovascular comorbidities1. Although RED-S occurs both in males and females, detection is easier in females, partly due to the fact that RED-S is likely to be exacerbated by symptoms related to menstrual dysfunction1,2. Low energy availability (LEA), one of the most important comorbidities of RED-S, represents a chronic negative energy balance resulting from a failure to satisfy real-time energy requirements. Consequently, LEA often results in prolonged states of undernourishment, a metabolic state with distinct biomarker profiles3. Functional hypothalamic amenorrhea (FHA), another common comorbidity of RED-S, is also associated with markedly altered metabolic profiles4.

LEA, as a negative energy balance state, can be used as a model to identify metabolic processes that are activated during energy deficiency. Identifying biomarkers associated with various cardiovascular, hormonal and anthropometric comorbidities in RED-S may improve our understanding of the biology of undernutrition and may highlight metabolic pathways which are important in states of metabolic imbalance.

Acute changes of metabolomic profiles in response to various exposures, such as exercise, and oral glucose- or lipid tolerance tests, have been investigated in the Human Metabolome Study in 15 healthy males5, but no such studies have been undertaken in undernourished females. Nonetheless, unfavourable lipid profiles have been

---

1Novo Nordisk Foundation Center for Protein Research, Translational Disease Systems Biology Group, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark. 2Department of Clinical Sciences, Genetic and Molecular Epidemiology Unit, Lund University, Skåne University Hospital Malmö, Malmö, Sweden. 3Steno Diabetes Center Copenhagen, Gentofte, Denmark. 4National Bioinformatics Infrastructure Sweden (NBIS), SciLifeLab, Lund, Sweden. 5Biosyntia ApS, Copenhagen, Denmark. 6Institute of Pharmaceutical Science, Faculty of Life Sciences and Medicine, King’s College London, London, United Kingdom. 7Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark. 8Endocrinological and Reproductive Unit, Department of Obstetrics and Gynecology, Herlev Hospital, Herlev, Denmark. 9Department of Health Sciences, Lund University, Lund, Sweden. 10These authors contributed equally: Søren Brunak and Åsa B. Tornberg. *email: soren.brunak@cpr.ku.dk; asa.tornberg@med.lu.se
LC-MS lipidomics. The 190 (N = 38 × 5) frozen blood serum samples were sent to the Metabolomics Laboratory at the Steno Diabetes Center Copenhagen (Gentofte, Denmark). The samples had not been thawed before this analysis. The laboratory procedures and the list of lipids used as standards and calibrants have been previously described. In brief, during sample preprocessing, 10 μl of 0.9% NaCl, 92 μl of chloroform:methanol (2:1, v/v) and 28 μl of a lipid standard solution (10 μg/ml) were added to 10 μl of serum samples. All samples and controls were mixed and placed on ice for a minimum of 30 minutes and subsequently centrifuged (3 min, 4 °C, 9400 g). Following centrifugation, 60 μl from the lower layer of each sample was transferred to a glass vial and mixed with 60 μl chloroform:methanol (2:1, v/v). Untargeted lipidomics analyses were performed by an ultra-high performance liquid chromatography quadrupole time-of-flight mass spectrometry (UHPLC-Q-TOF-MS) equipment by Agilent Technologies (Santa Clara, CA, USA). The raw lipidomic data was processed using MZmine 213, an open-source software for mass-spectrometry data processing. In brief, the processing steps included mass detection, chromatogram builder, chromatogram deconvolution, isotopic peak grouper, peak filter, join aligner, peak list row filter, gap filling and annotation to in-house library. In total, 201 lipids could be annotated using
the Steno Diabetes Center internal peak library based on m/z values and retention time. These metabolites represent a subset that we consider to be annotated with higher confidence and thus, are the main focus of our investigation. The rest of the lipids were annotated based on their m/z values using LIPID MAPS 15,16, a comprehensive and widely used online resource for lipidomics annotations. As compounds annotated using LIPID MAPS have lower confidence annotations, we exercised caution in presenting individual associations with these features in downstream analyses. After data processing and annotation, the available lipids (n = 1,060) were further filtered. Lipids with ambiguous annotations were removed: these included lipids used as standards (n = 17), lipids with no annotations (n = 167) and duplicated LIPID MAPS annotations (n = 242). The remaining named, unique lipids (n = 634) comprise the final lipidomics dataset used in the study.

Statistical analyses. Statistical analyses were performed using R v3.5.227 and Python 3.6.18. The median, mean and maximum feature missingness rates in the dataset were 1.1%, 3.7% and 45%, respectively. Missing data of the unscaled data was imputed using the missForest R package 28, an algorithm utilizing random forest to impute missing data in the data matrix based on observed datapoints.

The main data analysis can be separated into two main parts:

Analyses at fasting state. For the first part of the analysis, the dataset was restricted to the first timepoint (fasting status). Highly correlated features (Pearson’s |r| > 0.8) were removed, which resulted in 216 lipidomic features for further analysis. All lipid features were scaled and centered (mean = 0, standard deviation = 1). Linear regression (24 unscaled numeric outcomes) and logistic regression (2 categorical outcomes) models were constructed with the individual lipid features being the independent variables. In these models, age was added as a covariate. In total, 26 × 216 = 5,616 statistical tests were performed. P values from these tests were corrected using the Benjamini-Hochberg false discovery rate (FDR) method (α = 0.05)29. To assess whether P < 0.05 findings were overrepresented among results for a given trait (compared to the other 25 traits), Fisher’s exact tests were employed. Here, P < 0.002 (α = 0.05/26, adjusted using the Bonferroni method) were considered statistically significant. To evaluate the cumulative predictive utilities of the 216 lipids towards the 24 scaled numeric outcomes and the 2 categorical outcomes, linear and logistic regression was utilized in a leave-one-out cross validation (LOOCV) framework, which has been shown to be better suited for small datasets compared to regular k-fold cross-validation. In this framework, for each outcome, models are trained on 37 samples and tested on the 38th sample. This results in 38 test sets, each test set being a participant in the study. For numeric outcomes, prediction error, measured by mean absolute error (MAE), and explained variance (R²) were subsequently averaged between the 38 test sets. For categorical outcomes, prediction accuracy is reported (number of correct predictions/number of all predictions).

Analyses considering all timepoints. For the second part of the analysis, data at all five timepoints were considered. All 634 lipids were stratified into distinct clusters using the Compass hybrid method 30. First, the 634 lipid features were scaled and centered into z-scores. Subsequently, Self-Organizing Maps (SOM) 31 were employed to reduce dimensionality and cluster participants based on the lipidomics data at all five timepoints, separately. In this method, a set of interconnected neurons is trained to adapt to the original data set. The way nodes are connected is defined beforehand; here we tested a hexagonal (6 connections) grids of sizes 4 × 4, 5 × 5, 6 × 6 and 7 × 7. The training process consists in reducing the Euclidean distances between the model neurons and the data points in a process called competitive learning. For each data sample, distances to the model vectors are calculated, and the closest neuron, also called as Best Matching Unit, is selected as the “winner”. Next, the values of the “winning” neuron and its neighbour nodes are adjusted towards the sample. This process is iterated until the network approximates to the original data space. Once the training process is finished, it is possible to assign samples to their Best Matching Units, which can be considered as clusters. In this study, the most optimal grid, based on fitness, was a 4 × 4 neuron structure. The resulting neurons were aggregated into larger groups using the Partitioning Around Medoids algorithm, estimating the number of clusters based on the optimum average silhouette width of the SOM nodes. Associations between the resulting SOM clusters and the two categorical variables, LEA and FHA were assessed using chi-squared tests. Associations with P < 0.05 were considered statistically significant.

Lipid trajectory clustering was undertaken using the traj R package. Here, highly correlated features (Pearson’s |r| > 0.8) were removed, which resulted in 245 lipid features in analysis. Using traj, trajectory clustering for all individual lipid features were undertaken separately and for each iteration, 24 meta-features describing the lipid trajectories were extracted. These meta-features included trajectory properties such as feature range, mean over time, change, slope, standard deviation, maximum differences between timepoints, etc. The obtained meta-features were subsequently visualized using a heatmap.

Results
Participant characteristics. Descriptive statistics of various biomarkers in LEA study participants have been published before32, but these statistics described study populations with slight differences in size and constitution. Therefore, medians and interquartile ranges (IQR) for all continuous clinical biomarkers at fasting state for all participants eligible for this study (together and stratified by LEA and FHA status) are shown in Table 1. Amongst the participants, 12 individuals had LEA, while 26 had sufficient energy availability (EA). Furthermore, while 27 participants were diagnosed with FHA, 11 participants had eumenorrhea (healthy menstrual cycle).

Classification of lipidomic features. The 634 identified named lipid features were organized into four major categories. The first category, Glycerolipids, was comprised of diacylglycerols (DG) and triacylglycerols (TG). The second category, Glycerophospholipids, was comprised of lysophosphatidylcholines (LPC),
phosphatidylcholines (PC), phosphatidylethanolamines (PE), phosphatidylglycerols (PG), phosphatidylserines (PS) and phosphatidic acids (PA). The third category, Miscellaneous, demonstrated statistically significant associations with lipidomic features. We present spectral data for the 216 features including cholesteryl esters (CE), vitamins, and other named features.

Table 1. Characteristics of LEA participants at fasting status (N = 38). Abbreviations: ACTH - adrenocorticotropic hormone; BDNF - brain-derived neurotrophic factor; BMD - bone mineral density; BMI - body mass index; DHA - dehydroepiandrosterone; EA - energy availability; EUM - eumenorrhea; FHA - functional hypothalamic amenorrhea; FSH - follicle stimulating hormone; GH - growth hormone; HDL-C - high-density lipoprotein cholesterol; LDL-C - low-density lipoprotein cholesterol; LEA - low energy availability; LH - luteinizing hormone; SHBG - sex hormone binding globulin; TC - total cholesterol; TG - triglycerides; TSH - thyroid stimulating hormone. The values shown are median values and interquartile ranges.

| Trait                  | All (n = 26) | Sufficient EA (n = 26) | LEA (n = 12) | EUM (n = 11) | FHA (n = 27) |
|------------------------|-------------|------------------------|-------------|-------------|-------------|
| Age (years)            | 20.35 [19.21;21.37] | 20.19 [18.93;21.19] | 20.8 [20.21;21.52] | 20.8 [20.29;22.03] | 20.2 [18.85;21.04] |
| HDL-C (mmol/l)         | 1.8 [1.62;2.07]       | 1.79 [1.43;2.07]       | 1.99 [1.74;2.12]       | 1.76 [1.29;2.3]       | 1.86 [1.66;2.06]       |
| LDL-C (mmol/l)         | 2.3 [2.12;2.9]         | 2.25 [1.95;2.9]         | 2.3 [2.27;2.82]         | 2.3 [2.25;2.5]         | 2.3 [2.3;2.3]         |
| TC (mmol/l)            | 4.55 [4.15;5.1]        | 4.35 [3.82;5.07]        | 4.7 [4.45;5.12]        | 4.6 [4.4;4.9]        | 4.5 [4.13;5.3]        |
| TG (mmol/l)            | 0.67 [0.61;0.87]       | 0.66 [0.580;0.79]       | 0.68 [0.64;61]         | 0.7 [0.660;0.91]      | 0.66 [0.59;0.81]      |
| Androstendion (nmol/l) | 4.82 [3.38;5.63]       | 5.28 [3.26;6.35]       | 4.09 [3.59;6.59]       | 4.08 [3.06;6.74]     | 5.26 [3.52;6.15]     |
| DHAS (nmol/l)          | 3328 [2455;4045]       | 3300 [2675;5065]       | 3229 [2134;3797]       | 3020 [2176;8333]     | 3360 [2955;4240]     |
| Estrogen (ng/ml)       | 0.12 [0.09;0.16]       | 0.12 [0.09;0.15]       | 0.14 [0.07;0.17]       | 0.13 [0.12;0.18]     | 0.11 [0.07;0.16]     |
| Free testosterone (nmol/l) | 0.01 [0.01;0.02] | 0.02 [0.01;0.02] | 0.01 [0.01;0.02] | 0.01 [0.01;0.01] | 0.02 [0.01;0.02] |
| FSH (U/l)              | 6.5 [5.7;8.1]          | 6.5 [5.9;2.81]         | 6.8 [5.27;8.7]         | 6.5 [5.85;7.5]       | 6.5 [5.7;8.4]       |
| LH (U/l)               | 4.15 [2.72;6.1]        | 4.15 [2.35;6.7]        | 4.5 [3.22;5.8]        | 3.4 [2.95;15]        | 4.9 [2.75;0.5]       |
| Progerston (ng/ml)     | 1.23 [0.84;1.48]       | 1.14 [0.74;1.47]       | 1.28 [1.18;1.48]       | 1.32 [0.96;1.55]     | 1.18 [0.77;3.16]     |
| Prolactin (U/l)        | 0.18 [0.13;0.24]       | 0.15 [0.12;0.19]       | 0.22 [0.18;0.27]       | 0.22 [0.18;0.26]     | 0.15 [0.11;0.21]     |
| SHBG (nmol/l)          | 69 [60.88]             | 69 [59.87]             | 70 [63.89]             | 69 [56.83]            | 70 [61.97]            |
| Total testosterone (nmol/l) | 1.14 [0.82;1.34] | 1.17 [0.82;1.34] | 1.06 [0.82;1.3] | 0.81 [0.51;1.19] | 1.17 [0.91;1.46] |
| ACTH (µg/l)            | 22 [17.36]             | 21 [16.36]             | 24 [22.30]             | 15 [13.24]            | 23 [19.37]            |
| BDNF (µg/l)            | 117 [78.16]            | 124 [90.16]            | 111 [94.16]            | 138 [112.16]         | 107 [72.163]         |
| Cortisol (nmol/l)      | 470 [393.550]          | 478 [398.554]          | 458 [380.537]          | 424 [279.461]        | 510 [423.555]        |
| GH (ng/ml)             | 4.2 [2.46;4.96]        | 4.36 [2.46;4.14]       | 3.76 [2.7.4.7]         | 2.9 [2.01;5.77]       | 4.73 [2.74;1.67]       |
| Insulin (mIU/l)        | 3.25 [2.45;1.5]        | 3.7 [2.6.2.6.2]        | 2.7 [2.25;3.6]        | 2.7 [2.25;4.6]        | 3.3 [2.56;15]         |
| Leptin (µg/ml)         | 2942 [1621;3875]       | 2151 [397;3549]        | 3170 [2915;4190]       | 3119 [1727;4988]     | 2608 [794;3524]       |
| TSH (mIU/l)            | 2.16 [1.69;2.86]       | 2.23 [1.53.28]         | 2.1 [1.96;2.66]       | 2.05 [1.42;5.42]     | 2.18 [1.77;2.92]     |
| BMD (kg/cm²)          | 1.11 [1.05;1.14]       | 1.1 [1.05;1.14]       | 1.12 [1.05;1.1]       | 1.13 [1.09;1.13]      | 1.09 [1.05;1.14]     |
| BMI (kg/m²)            | 20.35 [19.21;21.37]    | 20.19 [18.83;21.19]    | 20.8 [20.21;21.52]    | 20.8 [20.29;22.03]  | 20.2 [18.85;21.04]  |
| Fat mass (%)           | 20.25 [18.07;22.31]    | 19.7 [17.47;21.37]    | 21.35 [19.87;23.7]    | 20.8 [19.92;4.45]   | 19.8 [17.25;21.75]   |

Association and prediction using lipidomics. First, linear regression (for the 24 continuous traits) and logistic regression models (for the two categorical traits, LEA and FHA) were employed to test associations with the correlation filtered lipidomic features (n = 216). In total, 5,616 (26 outcomes × 216 features) statistical tests were performed. After FDR correction, 20 statistically significant associations were detected, from which nine had annotations using the Steno Diabetes Center in-house library (Table 2) and eleven using LIPID MAPS (Supplementary Table 1). In this set of results, fat mass percentage, 10% of the variance of leptin levels and predicted BMI levels with a MAE of 0.4 standard deviations. For all other outcomes, the lipidomic features demonstrated poor predictive utilities with explained variances <5% and MAE > 0.5 standard deviations. The lipidomic features also demonstrated poor to mediocre predictive utilities for LEA (accuracy = 0.5) and FHA status (accuracy = 0.7).
Clustering and lipid trajectories. Self-organizing maps were utilized to cluster individuals based on their lipidomic profiles. No correlation filtering was utilized before this analysis, therefore all 634 lipidomic features were used. The resulting clusters (as categorical variables) were subsequently tested for associations with LEA and FHA status using chi-squared tests. In these analyses, LEA status and the clusters at fasting state (three clusters identified) were nominally statistically significantly associated ($P = 0.03$). LEA and FHA status were not associated with the SOM clusters at any other timepoints ($P > 0.05$). A heatmap of the 201 lipidomic features annotated using the Steno Diabetes Center in-house library (higher confidence annotations) at fasting state, together with individual FHA status, LEA status, and SOM clusters is shown in Figure 3.

Visual presentation of the lipidomics trajectories for 201 lipids annotated using the Steno Diabetes Center in-house library over the five timepoints for all participants, and stratified by LEA and FHA status, organized by lipid classes, are available at Supplementary Text 2. This browsable material serves as a resource to generate

| Outcome | Lipidomic feature | $\beta$ | SE  | $P$     | $P_{FDR}$ | Annotation type | m/z value | Retention time |
|---------|-------------------|--------|-----|---------|-----------|-----------------|-----------|---------------|
| Cortisol (mmol/l) | LPC(22:5) | 81.63 | 17.83 | 5.42E-05 | 0.035 | SDC in-house | 570.36 | 3.80 |
| Cortisol (mmol/l) | LPC(18:2) | 77.08 | 18.39 | 1.71E-04 | 0.048 | SDC in-house | 520.34 | 3.79 |
| Cortisol (mmol/l) | PC(42:6) | 77.36 | 18.35 | 1.60E-04 | 0.048 | SDC in-house | 862.63 | 7.11 |
| HDL-C (mmol/l) | SM(d38:1) | 0.233 | 0.041 | 1.56E-06 | 0.009 | SDC in-house | 759.64 | 7.82 |
| HDL-C (mmol/l) | SM(d32:1) | 0.200 | 0.045 | 9.29E-05 | 0.041 | SDC in-house | 675.54 | 6.35 |
| HDL-C (mmol/l) | PC(O-36:2) | 0.200 | 0.045 | 9.38E-05 | 0.041 | SDC in-house | 772.62 | 7.58 |
| TC (mmol/l) | SM(d33:1) | 0.488 | 0.111 | 9.41E-05 | 0.041 | SDC in-house | 689.56 | 6.63 |
| TC (mmol/l) | SM(d18:2/18:1) | 0.482 | 0.112 | 1.18E-04 | 0.047 | SDC in-house | 727.57 | 6.48 |
| TC (mmol/l) | CE(18:2) + Unknown CE(667.6219) | 0.474 | 0.113 | 1.62E-04 | 0.048 | SDC in-house | 1315.20 | 9.72 |

Table 2. Associations between lipidomic features with Steno Diabetes Center in-house library annotations and clinical biomarkers at fasting state ($N = 38$). Abbreviations: $\beta$ - effect sizes from linear regression models; CE - cholesteryl ester; HDL-C - high-density lipoprotein cholesterol; LPC - lysophosphatidylcholine; m/z - mass-to-charge ratio; PC - phosphatidylcholine; SDC - Steno Diabetes Center; SE - standard error; SM - sphingomyelin; TC - total cholesterol. $P$ values were calculated using linear regression models. $P_{FDR}$ values were adjusted using the Benjamini-Hochberg false discovery rate (FDR = 0.05). Annotation types are described in Materials and Methods.
hypotheses on candidate lipidomic features in future studies. The visual inspection of the stratified mean lipid trajectories showed interesting patterns for a number of features. Multiple triglycerides with long-chain fatty acids, for instance TG(47:1), TG(51:1), or TG(55:3), demonstrate a pattern reactive to the exercise interventions in participants with eumenorrhea, while participants with FHA demonstrate markedly different patterns, for instance a steady increase throughout the study day (Figure 4A). SM(d41:1), a sphingomyelin, demonstrates trajectories differing in magnitude between individuals with LEA and sufficient EA status (Figure 4B). Another example is PE(O-38:5) or PE(P-38:4), a phosphatidylethanolamine, where participants with eumenorrhea and with sufficient EA status show a trajectory reactive to the exercise interventions. Here, while individuals with FHA appear to demonstrate a blunted response to the interventions, individuals with LEA show a similar reactive trajectory, however mean levels are considerably lower compared to those with sufficient EA (Figure 4C,D).

**Discussion**

In this study, fasting lipidomic profiles and acute adaptations of lipidomic features to standardized exercise tests were investigated in female elite athletes at risk of RED-S. The focus of this investigation was: (i) to assess whether lipidomic features associate with clinical biomarkers at fasting state and (ii) test whether trajectories of lipidomic features could provide information on metabolic responses to intensive exercise.

Previous investigations showed that unhealthful lipid profiles associate with RED-S comorbidities such as menstrual dysfunction and BMD in female athletes. Here, we investigated whether the lipidome, which offers a higher resolution of lipid profiles compared to standard laboratory lipid assessments, would associate with a wide range of RED-S comorbidities. Such comorbidities included lipid, hormonal, anthropometric and energy-related clinical biomarkers. In total, 20 associations were detected at fasting and four traits showed overrepresentation of statistically significant associations. While the associations with standard laboratory lipids, such as TC, HDL-C and LDL-C cannot be considered surprising, it is noteworthy that cortisol levels positively associated with seven features and also showed a general overrepresentation of statistically significant findings. It has been shown that cortisol levels are chronically mildly elevated in elite athletes, and that FHA status is shown to be associated with higher levels of fasting serum cortisol levels. Thus, we hypothesize that the LPC and PC features associated with cortisol might serve as diagnostic proxies for a metabolically altered state resulting from RED-S, and are likely correlates of RED-S severity.

The study of acute and long-term trajectories of lipidomic profiles has the ability to reveal metabolic adaptations that would normally remain hidden, as standard clinical biomarkers cannot provide the data resolution achieved by various omics platforms. Acute changes of metabolomic profiles (lipids and amino acids) in response to various metabolic stressors have been tested in smaller studies with strict protocols, whereas long-term adaptations have been studied in larger epidemiological studies. These studies showcase that...
complex metabolic profiles are highly adaptive and demonstrate marked differences between various groups of individuals, for instance between males and females. The authors acknowledge that the individual lipidomic feature trajectories stratified by LEA and FHA status, due to the low sample size, are exploratory in nature. While no obvious clusters could be identified using the trajectory clustering algorithm, the presented individual lipidomic trajectories serve as a resource for future investigations which seek candidate biomarkers showing promising clinical or predictive utility, based on lipid trajectory differences between individuals stratified by menstrual dysfunction and EA. The visual inspection of all plots revealed a number of lipidomic features, belonging to various biological subclasses (e.g. LPC, PC, SM, TG), demonstrating tentative evidence for lipid trajectory differences between individuals stratified by FHA status. Lipid trajectory differences between individuals with sufficient EA and LEA were lower in numbers. These results tentatively indicate that, in this cohort of female elite athletes, FHA status is associated with an altered ability for metabolic adaptation to exercise interventions.

Our study has a number of important strengths. First, deep phenotyping was undertaken, i.e. a wide range of lipid, hormonal, anthropometric, and energy-related biomarkers were assessed together with lipidomic profiles. Second, this study is prospective in nature; the participants underwent a day-long protocol and lipidomic profiles were obtained at five different timepoints throughout the day, at fasting state, and before and after two exercise sessions. This study design and data collection presents a rare opportunity to study adaptive, acute changes in metabolic profiles. Third, a robust statistical framework used biomarkers at fasting state, along with the longitudinal aspect of the study, the lipid trajectories. Fourth, all participants adhered to a strict study protocol with standardized exercise interventions. Fifth, objective measurements of anthropometric and energy-related traits and reproductive function (e.g. the thorough classification of menstrual functions) were available. Sixth, we provide reference trajectories for 201 lipidomic features, which may provide interesting candidate features for future investigations.

The most obvious limitation of this project is the low sample size. The authors acknowledge that due to the small number of participants, findings from this study are exploratory in nature, and larger cohorts must be utilized for validation. However, as expressed above, due to challenges related to establishing studies with standardized prospective designs with high-frequency sampling and deep phenotyping, our sample size is not unusual. Indeed, previous studies that demonstrated the dynamic nature of metabolomic adaptations in response to prolonged fasting, cold stress, exercise, meal challenge, and oral glucose tolerance tests had lower sample sizes and lower number of metabolites assessed. To ameliorate the challenges posed by low sample size, we utilized statistical approaches appropriate for smaller samples sizes, such as presenting nonparametric test statistics and using LOOCV for prediction analyses. A high number of statistical tests were undertaken, which can result in increased type 1 error rates. We addressed this by applying multiple testing corrections, Benjamini-Hochberg FDR for lipidomics associations and the more conservative Bonferroni correction for other comparisons. We acknowledge that annotations obtained by LIPID MAPS yielded a number of metabolites that are unlikely to be present in human serum samples. Thus, when presenting our results, we strived to primarily present findings with higher confidence annotations obtained by our in-house library, all confirmed to be present in human blood.

Figure 4. Lipid trajectories of TG(55:3), SM(d41:1) and PE(O-38:5) or PE(P-38:4). The figures show mean trajectories (using locally estimated scatterplot smoothing) for all LEA participants (black line), participants with sufficient EA/eumenorrhea (red) and LEA/FHA (cyan). Panel (A) shows trajectories for TG(55:3) stratified by FHA status. Panel (B) shows trajectories for SM(d41:1) stratified by LEA status. Panels (C,D) show trajectories for PE(O-38:5) or PE(P-38:4) stratified by FHA and LEA status, respectively.
Last, we acknowledge that LEA is a highly specific cohort; while female elite athletes at risk for RED-S represent an interesting population to study, findings from this project are likely to have limited generalizability to different populations, for instance males at risk for RED-S or for non-athlete females.

In summary, we demonstrate associations between a number of lipidomic features and cortisol and standard laboratory lipids while in a fasting state. We present reference lipid trajectories for hundreds of lipidomic features, many of which demonstrate distinct patterns for participants with menstrual dysfunction or undernourishment.

Acknowledgements

The authors thank Birgitte Nergaard Roberts (Steno Diabetes Center Copenhagen) for the blood sample preprocessing. We thank Katrine Haahr (Herlev Hospital) for performing the gynecological examinations; Anders Johansson (Lund University) and Ulla Kjærulf-Hansen (Herlev Hospital) for taking the blood samples; Fiona Koivola (Lund University) for assisting during testing; Hanne Udengaard (Herlev Hospital) for logistic assistance; Inge Birgit Rasmussen and Søren Andresen (University of Copenhagen) for blood analysis. We highly appreciate the support of the Swedish and Danish National Sport Federations and Team Danmark. We thank Catherine Bjerre Collin for her feedback on the manuscript. T.V.V. is supported by the Novo Nordisk Foundation Postdoctoral Fellowship within Endocrinology/Metabolism at International Elite Research Environments via

Received: 25 September 2019; Accepted: 23 January 2020; Published online: 11 February 2020

References

1. Mountjoy, M. et al. IOC consensus statement on relative energy deficiency in sport (RED-S): 2018 update. Br. J. Sports Med. 52, 687–697, https://doi.org/10.1136/bjsports-2018-099193 (2018).
2. Loucks, A. B. Low energy availability in the marathon and other endurance sports. Sports Med. 37, 348–352, https://doi.org/10.2165/00007256-200737040-00019 (2007).
3. Melin, A. et al. Energy availability and the female athlete triad in elite endurance athletes. Scand. J. Med. Sci. Sports 25, 610–622, https://doi.org/10.1111/sms.12261 (2015).
4. Melin, A. K. et al. Impact of Menstrual Function on Hormonal Response to Repeated Bouts of Intense Exercise. Front. Physiol. 10, 1–8, https://doi.org/10.3389/fphys.2019.00942 (2019).
5. Krug, S. et al. The dynamic range of the human metabolome revealed by challenges. FASEB journal: Off. Publ. Federation Am. Societies Exp. Biol. 26, 2607–2619, https://doi.org/10.1096/fj.11-198093 (2012).
6. Soleimany, G. et al. Bone Mineral Changes and Cardiovascular Effects among Female Athletes with Chronic Menstrual Dysfunction. Asian J. Sports Med. 3, 53–58 (2012).
7. Rickenland, A., Eriksson, M. J., Schenck-Gustafsson, K. & Hirschberg, A. L. Amenorrhea in female athletes is associated with endothelial dysfunction and unfavorable lipid profile. J. Clin. Endocrinol. Metab. 90, 1354–1359, https://doi.org/10.1210/jc.2004- 1286 (2005).
8. Melin, A. et al. Low-energy density and high fiber intake are dietary concerns in female endurance athletes. Scand. J. Med. Sci. Sports 26, 1060–1071, https://doi.org/10.1111/sms.12516 (2016).
9. Fahrenholze, I. L. et al. Within-day energy deficiency and reproductive function in female endurance athletes. Scand. J. Med. Sci. Sports 28, 1139–1146, https://doi.org/10.1111/sms.13030 (2018).
10. Tornberg, A. B. et al. Reduced Neuromuscular Performance in Amenorrheic Elite Endurance Athletes. Med. Sci. sports Exerc. 49, 2478–2485, https://doi.org/10.1249/MSS.0b013e3181e53383 (2017).
11. Nattiv, A. et al. American College of Sports Medicine position stand. The female athlete triad. Med. Sci. sports Exerc. 39, 1867–1882, https://doi.org/10.1249/MSS.0b013e3181e53383 (2007).
12. O’Gorman, A. et al. Identification of a plasma signature of psychotic disorder in children and adolescents from the Avon Longitudinal Study of Parents and Children (ALSPAC) cohort. Transl. psychiatry 7, e1240, https://doi.org/10.1038/tp.2017.211 (2017).
13. Pluskal, T., Castillo, S., Villar-Briones, A. & Oresic, M. MZmine 2: modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. BMC Bioinforma. 11, 395, https://doi.org/10.1186/1471-2105-11-395 (2010).
14. Luukkonen, P. K. et al. Impaired hepatic lipid synthesis from polyunsaturated fatty acids in Tmem65E2 E167K variant carriers with NAFLD. J. hepatology 67, 128–136, https://doi.org/10.1016/j.jhep.2017.02.014 (2017).
15. Sud, M. et al. LMSD: LIPID MAPS structure database. Nucleic Acids Res. 35, D527–532, https://doi.org/10.1093/nar/gkl838 (2007).
16. Fahy, E., Sud, M., Cotter, D. & Subramaniam, S. LIPID MAPS online tools for lipid research. Nucleic Acids Res. 35, W606–612, https://doi.org/10.1093/nar/gkm524 (2007).
17. R: A language and environment for statistic computing (R Foundation for Statistic Computing, Vienna, Austria, 2015).
18. Python Software Foundation. Python Language Reference, version 3.6. Available at, http://www.python.org.
19. Stehkoven, D. J. & Buhlmann, P. MissForest–non-parametric missing value imputation for mixed-type data. Bioinforma. 28, 112–118, https://doi.org/10.1093/bioinformatics/btr597 (2012).
20. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Society. Ser. B 57(1), 125–133, (1995).
21. Krysiak-Baltyn, K. et al. Compass: a hybrid method for clinical and biobank data mining. J. Biomed. Inf. 47, 160–170, https://doi.org/10.1016/j.jbi.2013.10.007 (2014).
22. Kohonen, T. Essentials of the self-organizing map. Neural Netw. 37, 52–65, https://doi.org/10.1016/j.neunet.2012.09.018 (2013).
23. Hasin, Y., Seldin, M. & Lusis, A. Multi-omics approaches to disease. Genome Biol. 18, 83, https://doi.org/10.1186/s13059-017-1215-1 (2017).
24. Yan, J., Risacher, S. L., Shen, L. & Saykin, A. J. Network approaches to systems biology analysis of complex disease: integrative methods for multi-omics data. Brief. Bioinform. 19, 1370–1381, https://doi.org/10.1093/bib/bbx066 (2018).
25. Wopereis, S. et al. Metabolic profiling of the response to an oral glucose tolerance test detects subtle metabolic changes. PLoS one 4, e4525, https://doi.org/10.1371/journal.pone.0004525 (2009).
26. Winnike, J. H., Bubis, M. G., Watkins, P. B. & O’Connell, T. M. Effects of a prolonged standardized diet on normalizing the human metabolome. Am. J. Clin. Nutr. 90, 1496–1501, https://doi.org/10.3945/ajcn.2009.28234 (2009).
27. Grace, M. S. et al. Breaking up prolonged sitting alters the postprandial plasma lipidomic profile of adults with Type 2 Diabetes. The journal of clinical endocrinology and metabolism, https://doi.org/10.1210/jc.2016-3926 (2017).
28. Mielke, M. M. et al. Factors affecting longitudinal trajectories of plasma sphingomyelins: the Baltimore Longitudinal Study of Aging. Aging Cell 14, 112–121, https://doi.org/10.1111/ace.12275 (2015).
29. Darst, B. F., Kosic, R. L., Hogan, K. J., Johnson, S. C. & Engelman, C. D. Longitudinal plasma metabolomics of aging and sex. Aging 11, 1262–1282, https://doi.org/10.18632/aging.101837 (2019).
NNF16OC0020698, the Swedish Research Council (Strategic Research Area Exodiab) via Dnr 2009-1039, and the Swedish Foundation for Strategic Research via Dnr IRC15-0067. J.A.R.H. and S.B. are supported by the Novo Nordisk Foundation via NNF14CC0001 and NNF17OC0027594. The LEA study was funded by research grants from the Faculty of Science, University of Copenhagen, the World Village of Women Sports Foundation, and Arla Foods Ingredients. Open access funding provided by Lund University.

Author contributions
The authors' responsibilities were as follows: The authors' responsibilities were as follows: T.V.V. designed the study. T.V.V. performed the statistical analyses. T.V.V., A.A. and N.H.A. performed the processing and quality control of the lipidomics data. L.L.A., I.M.M. and N.H.A. performed the laboratory lipidomics assessments. T.V.V. and J.A.R.H. performed the Compass clustering. S.S. undertook the gynecological examinations in the LEA Study. S.B. and Å.B.T. provided supervision and guidance on methodology and aims. T.V.V. drafted the manuscript. All authors critically revised and approved the manuscript. T.V.V., S.B. and Å.B.T. have primary responsibility for final content.

Competing interests
S.B. has ownerships in Intomics A/S, Hoba Therapeutics Aps, Novo Nordisk A/S, Lundbeck A/S and managing board memberships in Proscion A/S and Intomics A/S. The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41598-020-59127-8.

Correspondence and requests for materials should be addressed to S.B. or Å.B.T.

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