Article

Comprehensive and Quantitative Profiling of the Human Sweat Submetabolome Using High-Performance Chemical Isotope Labeling LC-MS

Kevin Hooton, Wei Han, and Liang Li

Just Accepted

Comprehensive and Quantitative Profiling of the Human Sweat Submetabolome Using High-Performance Chemical Isotope Labeling LC-MS

Kevin Hooton, Wei Han and Liang Li*

Department of Chemistry, University of Alberta, Edmonton, Alberta T6G 2G2, Canada

* Corresponding author.   E-mail: Liang.Li@ualberta.ca.

Abstract

Human sweat can be noninvasively collected and used as a media for diagnosis of certain diseases as well as for drug detection. However, because of very low concentrations of endogenous metabolites present in sweat, metabolomic analysis of sweat with high coverage is difficult, making it less widely used for metabolomics research. In this work, a high-performance method for profiling the human sweat submetabolome based on chemical isotope labeling (CIL) LC-MS is reported. Sweat was collected using a gauze sponge style patch and extracted from the gauze by centrifugation and then derivatized using CIL. Differential ^{12}C- and ^{13}C-dansylation labeling was used to target the amine/phenol submetabolome. Because of large variations in the total amount of sweat metabolites in individual samples, sample amount normalization was first performed using LC-UV after dansylation. The ^{12}C-labeled individual sample was then mixed with an equal amount of ^{13}C-labeled pooled sample. The mixture was subjected to LC-MS analysis. Over 2707 unique metabolites were detected across 54 sweat samples from six individuals with an average of 2002±165 metabolites detected per sample from a total of 108 LC-MS runs. Using a dansyl standard library, we were able to identify 83 metabolites with high confidence; many of them have never been reported to be present in sweat. Using accurate mass search against human metabolome libraries, we putatively identified an additional 2411 metabolites. Uni- and multivariate analyses of these metabolites showed significant differences in the sweat submetabolomes between male and female, as well as between early and late exercise. These results demonstrate that the CIL LC-MS method described can be used to profile the human sweat submetabolome with high metabolomic coverage and high quantification accuracy to reveal metabolic differences in different sweat samples, thereby allowing the use of sweat as another human biofluid for comprehensive and quantitative metabolomics research.

Introduction

Sweat is a colourless, hypotonic biofluid produced by both eccrine and aprocrine sweat glands in the epidermis. Its main function is the thermoregulation of internal body temperature through evaporative cooling in response to heat or physical exercise. The main constituents of sweat include water, electrolytes (sodium, potassium, and chloride), urea, pyruvate, lactate, and amino acids; but other molecules such as proteins, peptides, drugs, and other xenobiotics can also be detected.[1] It has been shown that the eccrine sweat proteome is significantly different than the serum proteome, suggesting that sweat components are not simply a diffused fluid from the plasma, but may be representative of surrounding tissues and cellular processes.[2] Thus, sweat can serve as another important biofluid for omics-based biomarker discovery research. Since the majority of sweat constituents are small molecules (<1000 Da), it is a prime candidate for metabolomics applications.[3] Currently, human urine and blood, as well as cerebrospinal fluid (CSF) and saliva, are more commonly used for metabolomics. Compared to collecting these biofluids, sweat sampling is attractive because it is a noninvasive procedure that can be conveniently collected on regular intervals or on a continuous basis.[4] For example, wearable sensors for monitoring a few sweat components (glucose, pH, Na and K) have been recently reported.[5] Since there are a total of 1.6 to 4.0 million sweat glands located on the majority of the epidermis,[6] sweat has enormous potential to become the perfect source of biomarkers for imaging the body's surface.

Sweat is already used as a clinical sample in a test for sweat chloride concentration to aid in the diagnosis of cystic fibrosis.[7] Certain drugs in sweat are detectable by mass spectrometry (MS) and sweat drug testing provides an alternative or complementary means of detecting substance abuse.[8] Despite these applications, this biofluid has been overlooked as a biospecimen

in most clinical analysis due to low concentration of endogenous metabolites, and lack of

research relating pathological states to metabolite composition. However, with recent advances

in the development of sensitive analytical techniques for metabolite detection, metabolomic

profiling of sweat may provide a means of searching for potential biomarkers of health state

using sweat samples. Indeed, a few studies of profiling sweat metabolites using NMR,[9, 10] GC-

MS[11, 12] and electrospray ionization (ESI)-MS or LC-MS[13-15] have been reported. However, the

overall metabolite coverage with current techniques is low (i.e., tens of metabolites detectable).[3]

In a related work, GC-MS analysis of volatile organic compounds (VOC) collected from human

skin including sweat[16] from 200 individuals detected a combined list of 373 peaks.[17] Aside from

limited coverage, due to large variations of total metabolite concentrations in sweat, sample

amount normalization is a major challenge for quantitative metabolomic profiling of individual

sweat samples.

We report a sensitive LC-MS workflow for human sweat metabolomics with high

coverage. Our method is based on the high-performance chemical isotope labeling (CIL) LC-MS

platform for quantitative and in-depth submetabolome profiling.[18-21] High-performance CIL uses

rationally designed labeling reagents to modify the metabolite's chemical and physical properties

to such an extent that simultaneous improvement in separation, ionization and quantification can

be achieved in LC-MS. For example, ^{12}C- and ^{13}C-dansyl labeling LC-MS can be used for

relative quantification of the amine/phenol submetabolome with high sensitivity.[18] This labeling

technique has been applied to metabolomics studies using serum, CSF,[22] urine,[23] saliva,[24] and

human feces.[25] In a recent study comparing the number of metabolites detectable by nano-flow

LC-MS and micro-flow LC-MS, human urine and sweat droplets were used as testing samples.[26]

In this report, the entire workflow for human sweat metabolomics is described and the analytical

performance of this method is demonstrated in metabolomic profiling of sweat samples collected at different time points during exercise. It is shown that a large number of metabolites could be readily profiled using CIL LC-MS. Thus sweat can be used as an important biofluid for comprehensive and quantitative metabolomics research.

Experimental Section

Figure 1 illustrates the overall workflow of the high-performance CIL LC-MS platform used in this study.

Sweat Sample Collection and Extraction. A total of 54 sweat samples were collected from both men (n=3) and women (n=3) with ethics approval from the University of Alberta. Volunteers were instructed to refrain from consuming alcohol 24 hr prior, and coffee the day of sweat collection. Samples were obtained on three separate days from each volunteer. On each day, sweat was collected from the proximal forearm for ten min during three separate time periods of moderate to intense exercise using a patch method.[27] To collect the sweat, a gauze sponge sweat patch was constructed in house and utilized. The proximal forearm collection area was first cleansed with 70% isopropanol and distilled water, followed by the application of a 10-layer, 5 cm × 5 cm gauze pad (Safe Cross First Aid Ltd); no cleansing was done between time intervals. Gauze was covered with a Teflon sheet (13/1000 inch) previously cleansed by sonicating in a mixture of organic solvents composed of acetonitrile, methanol, isopropanol and water, and secured with medical tape. Commercial plastic wrap (Glad® Cling Wrap) was wrapped around the forearm to prevent lifting of the tape. This was found to be an optimal method for collecting pure sweat, as the use of the Teflon sheet shortens the time that sweat is exposed to the air during exercise, thereby minimizing environmental contamination. For method

blank, water spiked gauze was sandwiched between 2 pieces of Teflon and then attached with the fixation tape.

We note that the use of the occlusive patch for a short period of time during exercise does not cause skin irritation. This simple and inexpensive "classical" method works for this type of application. We could collect droplets of sweat directly; but it is not easy and convenient, as we wanted to collect sweat cumulated over a period time (10-20, 20-30 and 30-40 min) to measure the averaged metabolite level changes in these three periods. Other methods such as using semi-permeable patches were not explored in this study. One concern with using semi-permeable patches is that any atmospheric chemicals surrounding a patch in a gym or outdoor area might contaminate the sweat collected in the patch. However, using semi-permeable patches is likely to be a better option for collecting sweat over a long period of time, as it reduces skin irritation. In the future, we will study the use of this option for sweat collection for metabolomics.

Volunteers were asked to lightly jog around an indoor track for 40 min, and sweat was collected during 10-20 (T1), 20-30 (T2), and 30-40 (T3) min of exercise. Samples were kept on ice until extraction via centrifugation at 14,000 rpm for 10 min. Extracted sweat was stored at -80°C until dansylation labeling. A total of 54 sweat samples were collected (6 volunteers × 3 time intervals × 3 separate days). A diagram of the sample collection is depicted in Supplemental Figure S1.

Dansylation Labeling. Frozen sweat samples were thawed at 4°C, vortexed to dissolve precipitates, then centrifuged at 14000 rpm for 10 min before aliquoting 12.5 μL of sweat for labeling. We did not study how sample freezing could affect the sweat metabolome and could not find any report of such a study, although sample freezing is not an issue for other biospecimens such as urine and serum. Future research of examining the sample storage issue on

sweat metabolome is warranted. In this work, the labeling method was adapted from a protocol previously described.[8] Supplemental Note S1 provides the detailed information on dansylation labeling.

LC-UV Sample Normalization. Sample amount was normalized using a protocol previously described based on LC-UV measurement of the total concentration of dansyl labeled metabolite in a sample.[28] A Waters ACQUITY UPLC system with photodiode array (PDA) detector was used to quantify the amount of labeled metabolites in the sweat samples. Labeled sweat was diluted 3-fold, and 4 μL of sample was injected onto a Phenomenex Kinetex C18 column (50 × 2.1 mm, 1.7 μm particle size, 100 Å pore size) for a fast step-gradient run to elute out all labeled metabolites together using a high percentage of organic solvent (see Supplemental Note S1 for running conditions used). The total peak areas of the labeled metabolites measured at 338 nm were used for sample quantification.

LC-QTOF-MS. Labeled sweat samples were analyzed using a Bruker HD Impact QTOF mass spectrometer (Billerica, MA, USA) with electrospray ionization (ESI) linked to an Agilent 1100 series HPLC system (Palo Alto, CA, USA) with an Agilent eclipse plus C18 column (100 × 2.1 mm, 1.8 μm particle size). Supplemental Note S1 provides the experimental conditions used.

Data Processing and Statistical Analysis. After LC-QTOF-MS analysis, entire peak lists were exported from Bruker Data Analysis software with a signal to noise threshold of three. IsoMS software was used for peak-pair picking, peak-pair filtering, and peak-pair intensity ratio calculations.[29] The program eliminates false-positive peaks such as dimers and common adducts. All 108 peak pair lists were aligned to produce a CSV file highlighting peak pairs shared between samples, as well as their accurate mass, retention time, and peak pair ratios relative to the control. A zero-fill program developed in-house was used to fill in missing values in the CSV

8

file by searching the raw data file for missed peaks.[30] Peak pairs were reconstructed and their chromatographic peak ratios were determined using IsoMS-Quant (m/z tolerance = 5 ppm, retention time tolerance = 30 s).[31]

Multivariate analysis was conducted using SIMCA-P+ (Version 12.0) software. Volcano plots were constructed using Microsoft Excel and OriginPro 8.5 (OriginLab). To calculate the fold-change between groups (male vs. female and different exercise times), the average peak pair ratio of all the injections in one group was calculated, then divided by the average peak pair ratio of all the injections in the other group. The p-value was calculated using a student's t-test. Peak pairs with a significant fold-change had a fold change $\geqslant$ 1.5, or $\leqslant$ 0.67, with $p \leqslant$ 0.05. The multiple-testing-corrected p-value (q-value) was calculated using R and BioConductor (www.bioconductor.org).[32]

Metabolite Identification. Positive metabolite identification was performed based on mass and retention time match to the dansyl standard library containing 273 unique amines/phenols using DnsID.[33] In our previous publication,[33] we have already shown that the combined information of accurate mass and retention time is sufficient for positive metabolite identification and there is no need of performing or using MS/MS information. Putative metabolite identification was done based on accurate mass match to the metabolites in the human metabolome database (HMDB) (8,021 known human endogenous metabolites) and the Evidence-based Metabolome Library (EML) (375,809 predicted human metabolites with one reaction) using MyCompoundID.[34] The mass accuracy tolerance window was set at 5 ppm for database search.

Results and Discussion

9

Sample Amount Normalization. When comparing individual metabolites present in different samples, it is crucial to normalize the total sample amount of these samples before performing metabolomic profiling. Electrolyte concentration in sweat has been found to increase in the event of dehydration compared to euhydration,[35] and it is possible that the concentration of metabolites may also vary depending on sweat rate and hydration levels. Sample normalization has been considered to be a major challenge in clinical sweat analysis.[8] However, normalizing the total amount of labeled metabolites in sweat samples is possible since dansyl labeling allows for the quantification of labeled metabolites using LC-UV.[28]

We applied the dansylation LC-UV method to quantify the sweat metabolites and Figure 2A shows a representative LC-UV chromatogram. The average concentrations of labeled metabolites in the 10-20, 20-30 and 30-40 min sweat were determined to be 1.25±0.40 mM, 0.60±0.09 mM, and 0.38±0.07 mM, respectively. Thus, there was more than 2-fold increase in concentration of labeled metabolites in sweat collected during early exercise (10-20 min) compared to late exercise (20-30 and 30-40 min). Greater volumes of sweat were obtained in samples collected from later exercise than earlier exercise. Sweat solute concentration was reported to be influenced by sweat rate,[36] which explains why the greater yield of sweat collected during late exercise was more dilute than sweat collected during early exercise. There was also larger variance in labeled metabolite concentration for the 10-20 min samples, compared to the 20-30 and 30-40 min samples (see Figure 2B). The larger variance in the 10-20 min sweat could be explained by differences in the fatigue of volunteers in early stages of exercise. During 10-20 min of exercise, some volunteers might experience fatigue more than others, resulting in larger amounts of sweating and more dilute sweat. In later stages of exercise, all volunteers were experiencing heavy fatigue and were sweating at a relatively more consistent rate, resulting in a

plateauing of sweat concentration. The large concentration differences found in sweat samples strongly suggest that sample amount normalization is required prior to quantitative metabolomic analysis. Note that, compared to a sample, the peak area of the blank is relatively small (i.e., 3% - 7%) (see Figures 2A and 2B) and thus during sample normalization the peak area of a blank was not subtracted.

Metabolite Detection. In order to demonstrate the performance of the developed method for comparative sweat metabolomics, we analyzed the amine/phenol submetabolome of sweat collected from 6 volunteers at three different days with three time points during exercise in each day. In total, 108 ^{12}C-/^{13}C-mixtures were analyzed from duplicate ^{12}C-labeling experiments of 54 sweat samples referencing to ^{13}C-labeled sweat pool. Amount of labeled metabolites injected onto LC-MS can have an effect on the number of metabolites detected in a sample. Therefore, injection amount was first optimized by injecting increasing volumes of a sweat sample with a known concentration of labeled metabolites that was determined using LC-UV. It was found that injecting 20 nmol of labeled metabolites generated the maximum number of peak pairs without sample carryover into subsequent injections; the sample amount was determined from the concentration of labeled samples measured by LC-UV. Thus, all sweat samples were analyzed using LC-MS with 20 nmol injection for each run. Figure 2C shows a representative base-peak ion chromatogram of dansyl labeled sweat.

From the 108 LC-MS runs, a total of 3130 different peak pairs were detected with an average of 2002±165 peak pairs per run. In order to differentiate the peak pairs resulting from the sweat samples and the background peak pairs resulting from the method, a method blank was studied. A new gauze pad of the same style was spiked with a small volume of water, followed by extracting and labeling the water using the same protocol as sweat labeling. The same volume

of labeled blank sample (30 µL) as the labeled sample was injected onto LC-MS for analysis. A total of 423 peak pairs were shared between the sweat samples and the method blank (see Supplemental Table S1 for the list). While it is possible that some of the 423 peak pairs from the blank may belong to the same compounds as those of sweat metabolites. We conservatively left all of them out and only counted the remaining pairs, i.e., 2707 unique pairs, as the sweat metabolites.

We note that the process of sweat collection can introduce background chemicals from multiple sources including chemical residuals remaining on the skin after cleaning, the gauze pad and extraction solvents. We have conducted a separate investigation of the peak pairs detected from a gauze pad without sweat, sweat collected in a gauze pad and a sweat droplet itself. Sweat collection was conducted on forehead, instead of forearm, as sweat droplet can be readily collected from forehead after exercise. The results are described in Supplemental Note S2, from which we conclude that the metabolites detected in sweat samples collected in a gauze pad were mainly from sweat (>90%).

Human Sweat Submetabolome. To identify the 2707 peak pairs unique to sweat, a dansyl-library search based on accurate mass and retention time was conducted against 315 entries from 273 unique dansylated metabolite standards. We were able to identify 83 metabolites with high confidence (see Supplemental Table S2A). Some of these metabolites such as amino acids are the same as those identified in sweat using other techniques.[9-15] We also searched the human metabolites in the HMDB library and the predicted human metabolites with one metabolic reaction in the MCID library using accurate masses of the peak pairs found. We matched 1025 metabolites to HMDB (Supplemental Table S2B) and 1386 metabolites to MCID (Supplemental Table S2C). Thus, out of the 2707 peak pairs detected, we could match a total of

2494 metabolites (92%). To our knowledge, this is the largest human sweat submetabolome ever reported. Our results indicate that the sweat metabolome is much more diverse than once thought, and its constituents may extend past the categories of electrolytes, amino acids, peptides, urea, and pyruvate.[1]

In all LC-MS runs, amino acids generated the strongest peak intensities compared to other metabolites (see Figure 2C). All common amino acids except histidine and a large number of dipeptides were also identified in sweat using dansyl-library search. Histidine was identified putatively using HMDB. Although lactate and urea were reported to have high abundance in sweat,[1] they could not be labeled by dansyl chloride due to their lack of primary or secondary amine or phenol group, and thus were not detected in our analysis. Note that the same workflow shown in Figure 1 should be applicable to other high-performance CIL experiments targeting different chemical-group-based submetabolomes. For example, DmPA has been shown to be useful for profiling the carboxyl acid submetabolome with high coverage.[19] Thus, we envisage that using multiple labeling of sweat will increase the metabolomic coverage even further.

Quantitative Metabolomics. While having high metabolomic coverage is certainly important to ensure that a large number of metabolites are examined, being able to quantify their changes in comparative samples with high accuracy and precision is also critical, particularly for revealing subtle differences. In the CIL LC-MS method, a pooled sample from mixing aliquots of individual sweat samples was labeled with ^{13}C-dansylation, which served as a global internal standard or control for analyzing all the ^{12}C-labeled individual samples (see Figure 1). Using the pooled sample as a reference mitigated the matrix effect and ion suppression effect on metabolite detection,[37] as a ^{12}C-labeled metabolite from a sample and its corresponding ^{13}C-labeled counterpart from the pool experienced the same labeling, separation, ionization and detection

processes with constant efficiencies. Thus, the peak ratio of the ^{12}C-/^{13}C-labeled metabolite shown as a peak pair in MS reflected the relative concentration of the metabolite in the sample vs. that in the pool. Since all the individual samples were spiked with the same pooled sample, the peak ratio values of a given metabolite found in all the samples provided the measurement of concentration changes of the metabolites in these samples.

Another important issue in quantitative metabolomics is related to the missing data in the metabolite-intensity table containing sample/metabolite IDs and their corresponding metabolite peak ratios. Missing data can affect the performance of multivariate analysis including Principle Component Analysis (PCA) and Partial Least Squares Discriminant Analysis (PLS-DA). In our work, we chose a stringent criterion to retain the metabolites for statistical analysis: metabolites must have peak ratio values in at least 80% of the samples in a comparative group. With this criterion, a total of 1262 peak pairs were used for multivariate analysis in comparing the sweat samples collected from males and females, and 1387 peak pairs were used to compare the 3 exercise intervals. When comparing sweat samples collected between individuals, metabolites must be present in 80% of all the injections. These three comparisons are described below to illustrate that the CIL LC-MS method is useful for quantitative sweat metabolomics.

Metabolomic Comparison of Individuals. Multivariate analysis was performed to compare time-dependent metabolomic changes in a series of samples collected from six individuals; three samples per day for three different days were collected from each individual. Figure 3 shows the PLS-DA plot of the metabolomic data set. Overall, higher inter-individual variability in the sweat metabolome was observed between subjects, which is consistent with previous studies involving sweat.[10] Sweat collected from individuals cluster together on the plot with $R^2X = 0.869$, $R^2Y = 0.961$ and $Q^2 = 0.931$, despite sample collection on 3 separate days

without any diet control. The model was validated using a permutations test (see Supplementary Figure S2A). This clustering indicates that an individual's genetics and environment (i.e., microbiome) are stronger determinants than diet on the sweat metabolome profile. Although an extensive study using more subjects and collecting more samples over a larger time frame is still needed to confirm this finding, this finding is not surprising considering that, even in plasma and urine collected from individuals with standardized meals, large inter-individual variability of the metabolomes has been observed.[38]

Metabolomic Comparison of Male and Female. There have been a number of studies investigating the effects of gender on the urine and plasma metabolome using either human or mouse models.[39-43] Studying the effects of gender, along with age, ethnicity and environment, on the human metabolome can assist in our understanding of the normal physiology of healthy individuals and help eliminate artifacts when conducting disease biomarker discovery investigations. We applied dansylation LC-MS to sweat samples collected from male and female individuals in order to determine whether this method is sufficiently sensitive and quantitative to detect metabolomic differences between genders. Figure 4A shows the PCA plot, and a clear separation between male and female sweat is visible on the second principle component. A separation is also observed mainly on the first principle component of the PLS-DA plot shown in Figure 4B, with $R^2X = 0.647$, $R^2Y = 0.973$ and $Q^2 = 0.960$. The model was validated using permutations test (see Supplemental Figure S2B).

The same data set was also subjected to univariate analysis to determine the significant metabolites that contribute to the separation between male and female sweat. In Figure 4C, a volcano plot highlights the significant metabolites with a fold-change of ≥ 1.5 or ≤ 0.67, and a p-value ≤ 0.05. In total, 161 peak pairs had an increase in concentration and 142 peak pairs had a

decrease in concentration in females compared to males. Among the 303 significant peak pairs, 14 of them were positively identified using the dansyl standard library, and 276 were putatively identified using accurate mass search against the HMDB and MCID libraries (see Supplemental Table S3A-C for the list). To gauge the false discovery rate, q-values were calculated for 1262 differentiating metabolites (see Supplemental Table S4). Using a cut-off of $p < 0.05$, before multiple-testing-correction, we expect a total number of 63 false positives with a $p < 0.05$. There should be far less false positives in the volcano plot since we also have the fold-change filter to determine the significant metabolites. The corresponding q-value for $p = 0.049907$ was 0.012406, meaning that 1.24% of metabolites with $p < 0.05$ were likely false positives. Thus the false discovery rate in this comparison is low.

Because there is no report of quantitative sweat metabolomics with this high coverage, direct comparison of our results to other sweat analyses could not be made. However, some of the metabolic changes found in our work are worth noting to indicate that sweat metabolomics is a largely untapped research field and our method allows investigating the sweat metabolome for biological and biomarker research. For example, previous research has shown an increase in hydroxyphenyllactic acid in male serum,[43] which is consistent with what we observed in male sweat. The L-form of this compound is found in higher concentrations in the urine and CSF of patients exhibiting phenylketonuria and tyrosinemia, while the D-form is only produced by bacteria and is found in patients with abnormal gut microbiota.[44] To our knowledge, this is the first instance of hydroxyphenyllactic acid being positively identified in human sweat, and this finding may encourage future work on discovering non-invasive biomarkers in sweat to aid in diagnosing these diseases. While some of the metabolites can be readily measured in blood test (e.g., for determining phenylketonuria), one could use sweat as an alternative media for

measuring these metabolite if the levels of these metabolites need to be monitored over a period of time. Multiple sample collections is in favor of sweat, as it is more convenient and less stress and risk to collect sweat than blood ,

Another example is on the observation of a 1.57-fold change for 6-hydroxynicotinic acid whose identify was confirmed using the dansyl standard library in female sweat compared to male sweat. This compound has been previously identified in rat and human urine after oral administration of nicotinic acid.[45,46] The metabolite is supposedly a downstream metabolite of nicotinic acid metabolism in *Pseudomonas* aeruginosa, and is used to identify *p. aeruginosa* in urinary tract infections.[47] *P. aeruginosa* is common bacteria of the natural human microbiome, and these results may be a reflection of female microbial flora compared to male. Although it is well known that the gut microbiome strongly influences the host's metabolic phenotype,[48, 49] to our knowledge, no studies have observed a significant gender difference in *p. aeruginosa* abundance. Future research in this direction, including direct analysis of microbes to determine the extent of microbe metabolite contribution to the human sweat metabolome, is warranted.

A third example is on the decreased levels of theophylline found in female sweat compared to male. Theophylline whose identify was confirmed using the dansyl standard library is a metabolic by-product of caffeine, which accounts for $3.7\pm1.3\%$ of caffeine demethylation.[50] It can also be demethylated by Cyp1A2, an enzyme with higher activity in men than in women.[51] The decreased level of theophylline in females may indicate lower levels of enzyme activities in females compared to males. It might also be possible that the females in this study consumed less caffeinated beverages by chance than the males. A large population-based study of sweat, in conjunction with other biospecimens such as blood, should be able to answer these questions.

The above results show that sweat metabolome differences between genders can be observed using dansylation LC-MS. Future work of using a larger sample size is needed to investigate and compare the sweat metabolomes of healthy individuals as baseline information for disease biomarker discovery research.

Metabolomic Comparison of Different Exercise Time. We envisage that quantitative sweat metabolomics with high coverage will play an important role in studying the relationship between exercise and human health; sweat can be readily collected during exercise. We were curious as to whether our method can be used to differentiate sweat collected during early exercise and late exercise. We thus compared sweat collected during 10-20 min (T1, early), 20-30 min (T2, middle), and 30-40 min (T3, late) of exercise and the results are shown in Figure 5. The PCA plot (Figure 5A) shows a gradual separation from early, middle and late exercise. The PLS-DA plot (Figure 5B) shows a clear separation between the early and late exercise, but the separation between the middle and late exercise is not as clear ($R^2X = 0.865$, $R^2Y = 0.978$, $Q^2 = 0.701$). The model was validated using a permutations test (see Supplemental Figure S2C).

Univariate analyses were also performed on the same data set. Volcano plots (Figure 6) compare the metabolite composition of sweat collected during different periods of exercise. Significant metabolites with a fold change of ≥ 1.5 or ≤ 0.67, and p-value ≤ 0.05 are highlighted. When comparing the composition of T1 vs. T2, 160 metabolites had an increase in concentration and 99 had a decrease in concentration in T2. When comparing the composition of T2 vs. T3, only 31 metabolites had a decrease in concentration in T3. When comparing the composition of T1 vs. T3, 299 metabolites had an increase in concentration and 411 had a decrease in concentration in T3. Of the 717 significant metabolites detected, 99% of them were detected when comparing T1 vs. T3. To gauge the false discovery rate, q-values of 1387 metabolites were

calculated (see Supplemental Table S5). Using a cut-off of $p < 0.05$, before multiple-testing-correction, we expect a total number of 69 false positives with a $p < 0.05$. Like the male vs. female sweat comparison, there should be far less false positives that we call significant after applying the fold-change filter. The corresponding q-value for $p < 0.047831$ was 0.006568, meaning that 0.66% of metabolites with $p < 0.05$ were likely false positives. Thus the false discovery rate in this comparison is also low.

Taken together, the results of both multivariate and univariate analyses indicate that larger metabolic changes are occurring between early exercise and late exercise, and smaller, indistinguishable changes in sweat composition occur once extreme fatigue has already set in. These metabolomic changes appear to be well correlated to the shift of aerobic metabolism to anaerobic metabolism during exercise.[52]. Because our technique can quantify many metabolite changes, it should be possible to design future experiments to study how this aerobic-to-anaerobic translation affects the overall metabolism in a time resolution that cannot be done using blood. For example, we could collect sweat samples frequently (e.g., by minutes) to observe the sweat metabolite changes with fast turn-over rates. One potential application of frequent sampling of sweat may be to monitor the metabolomic composition changes as a function of exercise time as a way of gauging healthy state of an individual.

Among the 710 significant metabolites separating T1 and T3, we positively identified 25 metabolites using the dansyl standard library and putatively identified 644 metabolites using HMDB and MCID database search (see Supplemental Table S6A-C for the list). The majority of the 25 identified metabolites with a significant fold-change after prolonged exercise are amino acids and dipeptides, and they seem to show a decrease in concentration during late exercise. Our findings are consistent with other studies that report a decrease in overall amino acid

concentration in the plasma after prolonged periods of exercise, mainly due to their catabolism to produce energy during anaerobic metabolism.[53, 54]

One of our interesting findings is related to imidazoleacetic acid, a common metabolite in the brain that is thought to be derived from the oxidation of histamine. This compound was found to have a decrease in concentration in T3 (late exercise) compared to T1 (early exercise). This observation may suggest a decrease in histamine metabolism or decrease in histamine release in late exercise under heavy fatigue. Imidazoleacetic acid has been shown to have neurophysiological, cardiovascular, and behavioral effects. While its mechanism of action and receptor site has been disputed, it has been repeatedly shown that introducing this compound into animal models causes a dose-dependent reduction in blood pressure.[55] It is possible that the body works to decrease imidazoleacetic acid abundance during late exercise in order to regulate blood pressure. Histamine was also detected with a fold-change of 0.50 in T3 compared to T1. In future studies, we will compare the sweat metabolome to those of other biofluids directly within the same individuals in order to gauge their differences.

Conclusions

We have developed a high-performance CIL LC-MS method for quantitative human sweat metabolome profiling with high coverage. Using dansylation LC-MS targeting the amine/phenol submetabolome, a total of 2707 unique metabolites were detected across 54 sweat samples from six individuals with an average of 2002±165 metabolites detected per sample from 108 LC-MS runs. We have detected many metabolites that have never been associated with human sweat, indicating that the human sweat metabolome is much more complex than previously thought. Because our technique is able to detect many metabolites, it opens the venue

20

for future research in designing biological or clinical experiments, including population-based study, instead of a limited number of subjects used in this study, to investigate their significances for biological or clinical applications. We conclude that, using CIL LC-MS, sweat can be used as another important human biofluid for in-depth and quantitative metabolomics research. Future work on improving detection sensitivity and sweat collection method may allow metabolomic monitoring of sweat using a volume of sample at a short interval.

Acknowledgements

Disclosure

The authors declare no competing financial interest.

Supporting Information

Supporting Information Available: Figures S1-S2, Note S1 on experiments and Note S2 on controls, Tables S1-S6 listing identification results and q-values. This material is available free of charge via the Internet at http://pubs.acs.org

References

(1)   Sato, K.; Kang, W. H.; Saga, K.; Sato, K. T. *J Am Acad Dermatol* **1989**, *20*, 537-563.

(2)   Raiszadeh, M. M.; Ross, M. M.; Russo, P. S.; Schaepper, M. A.; Zhou, W.; Deng, J.; Ng, D.; Dickson, A.; Dickson, C.; Strom, M., et al. *J Proteome Res* **2012**, *11*, 2127-2139.

(3)   Mena-Bravo, A.; Luque de Castro, M. D. *J Pharm Biomed Anal* **2014**, *90*, 139-147.

(4)   Bandodkar, A. J.; Wang, J. *Trends Biotechnol* **2015**, *32*, 363-371.

(5)   Gao, W.; Emaminejad, S.; Nyein, H. Y. Y.; Challa, S.; Chen, K.; Peck, A.; Fahad, H. M.; Ota, H.; Shiraki, H.; Kiriya, D., et al. *Nature* **2016**, *529*, 509-514.

(6)   Farrell, P. M.; Rosenstein, B. J.; White, T. B.; Accurso, F. J.; Castellani, C.; Cutting, G. R.; Durie, P. R.; Legrys, V. A.; Massie, J.; Parad, R. B., et al. *J Pediatr* **2008**, *153*, S4-S14.

(7)   Collie, J. T. B.; Massie, R. J.; Jones, O. A. H.; LeGrys, V. A.; Greaves, R. F. *Pediatr Pulmonol* **2014**, *49*, 106-117.

(8)   De Giovanni, N.; Fucci, N. *Curr Med Chem* **2013**, *20*, 545-561.

(9)   Kutyshenko, V. P.; Molchanov, M.; Beskaravayny, P.; Uversky, V. N.; Timchenko, M. A. *PLoS One* **2011**, *6*, 1-9.

(10)   Harker, M.; Coulson, H.; Fairweather, I.; Taylor, D.; Daykin, C. A. *Metabolomics* **2006**, *2*, 105-112.

(11)   Mark, H.; Harding, C. R. *Int J Cosmet Sci.* **2013**, *35*, 163-168.

(12)   Nunome, Y.; Tsuda, T.; Kitagawa, K. *Anal Sci* **2010**, *26*, 917-919.

(13)   Calderon-Santiago, M.; Priego-Capote, F.; Jurado-Gamez, B.; de Castro, M. D. L. *J Chromatogr A* **2014**, *1333*, 70-78.

(14)   Calderon-Santiago, M.; Priego-Capote, F.; Turck, N.; Robin, X.; Jurado-Gamez, B.; Sanchez, J. C.; de Castro, M. D. L. *Anal Bioanal Chem* **2015**, *407*, 5381-5392.

(15)   Dutkiewicz, E. P.; Lin, J. D.; Tseng, T. W.; Wang, Y. S.; Urban, P. L. *Anal Chem* **2014**, *86*, 2337-2344.

(16)   Soini, H. A.; Bruce, K. E.; Klouckova, I.; Brereton, R. G.; Penn, D. J.; Novotny, M. V. *Anal Chem* **2006**, *78*, 7161-7168.

(17)   Xu, Y.; Gong, F.; Dixon, S. J.; Brereton, R. G.; Soini, H. A.; Novotny, M. V.; Oberzaucher, E.; Grammer, K.; Penn, D. J. *Anal Chem* **2007**, *79*, 5633-5641.

(18)   Guo, K.; Li, L. *Anal Chem* **2009**, *81*, 3919-3932.

(19)   Guo, K.; Li, L. *Anal Chem* **2010**, *82*, 8789-8793.

(20)   Zhou, R.; Huan, T.; Li, L. *Anal Chim Acta* **2015**, *881*, 107-116.

(21)   Peng, J.; Li, L. *Anal Chim Acta* **2013**, *803*, 97-105.

(22)   Guo, K.; Bamforth, F.; Li, L. *J Am Soc Mass Spectrom* **2011**, *22*, 339-347.

(23)   Peng, J.; Chen, Y. T.; Chen, C. L.; Li, L. *Anal Chem* **2014**, *86*, 6540-6547.

(24)   Zheng, J.; Dixon, R. A.; Li, L. *Anal Chem* **2012**, *84*, 10802-10811.

(25)   Xu, W.; Chen, D.; Wang, N.; Zhang, T.; Zhou, R.; Huan, T.; Lu, Y.; Su, X.; Xie, Q.; Li, L., et al. *Anal Chem* **2015**, *87*, 829-836.

(26)   Li, Z. D.; Tatlay, J.; Li, L. *Anal Chem* **2015**, *87*, 11468-11474.

(27)   Verde, T.; Shephard, R. J.; Corey, P.; Moore, R. *J Appl Physiol* **1982**, *53*, 1540-1545.

(28)   Wu, Y.; Li, L. *Anal Chem* **2012**, *84*, 10723-10731.

(29)   Zhou, R.; Tseng, C. L.; Huan, T.; Li, L. *Anal Chem* **2014**, *86*, 4675-4679.

(30)   Huan, T.; Li, L. *Anal Chem* **2015**, *87*, 1306-1313.

(31)   Huan, T.; Li, L. *Anal Chem* **2015**, *87*, 7011-7016.

(32)   Gentleman, R. C.; Carey, V. J.; Bates, D. M.; Bolstad, B.; Dettling, M.; Dudoit, S.; Ellis, B.; Gautier, L.; Ge, Y. C.; Gentry, J., et al. *Genome Biol* **2004**, *5*, 16.

(33)   Huan, T.; Wu, Y. M.; Tang, C. Q.; Lin, G. H.; Li, L. *Anal Chem* **2015**, *87*, 9838-9845.

(34)   Li, L.; Li, R. H.; Zhou, J. J.; Zuniga, A.; Stanislaus, A. E.; Wu, Y. M.; Huan, T.; Zheng, J. M.; Shi, Y.; Wishart, D. S., et al. *Anal Chem* **2013**, *85*, 3401-3408.

(35)   McLean, C.; Graham, T. E. *J Appl Physiol (1985)* **2002**, *93*, 1471-1478.

(36)   Mangos, J. *Am J Physiol* **1973**, *224*, 1235-1240.

(37)   Han, W.; Li, L. *Metabolomics* **2015**, *11*, 1733-1742.

(38)   Kim, K.; Mall, C.; Taylor, S. L.; Hitchcock, S.; Zhang, C.; Wettersten, H. I.; Jones, A. D.; Chapman, A.; Weiss, R. H. *PLoS One* **2014**, *9*, e86223.

(39)   Slupsky, C. M.; Rankin, K. N.; Wagner, J.; Fu, H.; Chang, D.; Weljie, A. M.; Saude, E. J.; Lix, B.; Adamko, D. J.; Shah, S., et al. *Anal Chem* **2007**, *79*, 6995-7004.

(40)   Kochhar, S.; Jacobs, D. M.; Ramadan, Z.; Berruex, F.; Fuerholz, A.; Fay, L. B. *Anal Biochem* **2006**, *352*, 274-281.

(41)   Hodson, M. P.; Dear, G. J.; Roberts, A. D.; Haylock, C. L.; Ball, R. J.; Plumb, R. S.; Stumpf, C. L.; Griffin, J. L.; Haselden, J. N. *Anal Biochem* **2007**, *362*, 182-192.

(42)   Plumb, R.; Granger, J.; Stumpf, C.; Wilson, I. D.; Evans, J. A.; Lenz, E. M. *Analyst* **2003**, *128*, 819-823.

(43)   Dunn, W. B.; Lin, W.; Broadhurst, D.; Begley, P.; Brown, M.; Zelena, E.; Vaughan, A. A.; Halsall, A.; Harding, N.; Knowles, J. D., et al. *Metabolomics* **2015**, *11*, 9-26.

(44)   Spaapen, L. J.; Ketting, D.; Wadman, S. K.; Bruinvis, L.; Duran, M. *J Inherit Metab Dis* **1987**, *10*, 383-390.

(45)   Iwaki, M.; Murakami, E.; Kikuchi, M.; Wada, A.; Ogiso, T.; Oda, Y.; Kubo, K.; Kakehi, K. *J Chromatogr B Biomed Sci Appl* **1998**, *716*, 335-342.

(46)   Gronwald, W.; Klein, M. S.; Zeltner, R.; Schulze, B. D.; Reinhold, S. W.; Deutschmann, M.; Immervoll, A. K.; Boger, C. A.; Banas, B.; Eckardt, K. U., et al. *Kidney Int* **2011**, *79*, 1244-1253.

(47)   Gupta, A.; Dwivedi, M.; Nagana Gowda, G. A.; Ayyagari, A.; Mahdi, A. A.; Bhandari, M.; Khetrapal, C. L. *NMR Biomed* **2005**, *18*, 293-299.

(48)   Li, M.; Wang, B.; Zhang, M.; Rantalainen, M.; Wang, S.; Zhou, H.; Zhang, Y.; Shen, J.; Pang, X.; Zhang, M., et al. *Proc Natl Acad Sci U S A* **2008**, *105*, 2117-2122.

(49)   Nicholson, J. K.; Holmes, E.; Wilson, I. D. *Nat Rev Microbiol* **2005**, *3*, 431-438.

(50)   Lelo, A.; Miners, J. O.; Robson, R. A.; Birkett, D. J. *Br J Clin Pharmacol* **1986**, *22*, 183-186.

(51)   Ou-Yang, D. S.; Huang, S. L.; Wang, W.; Xie, H. G.; Xu, Z. H.; Shu, Y.; Zhou, H. H. *Br J Clin Pharmacol* **2000**, *49*, 145-151.

(52)   De Feo, P.; Di Loreto, C.; Lucidi, P.; Murdolo, G.; Parlanti, N.; De Cicco, A.; Piccioni, F.; Santeusanio, F. *J Endocrinol Invest* **2003**, *26*, 851-854.

(53)   Haralambie, G.; Berg, A. *Eur J Appl Physiol* **1976**, *36*, 39-48.

(54)   Refsum, H. E.; Gjessing, L. R.; Strømme, S. B. *Scand J Clin Lab Invest* **1979**, *39*, 407-413.

(55)   Tunnicliff, G. *Gen Pharmacol* **1998**, *31*, 503-509.

Figure Captions

Figure 1.   Workflow for sweat metabolomics using a high-performance CIL LC-MS method.

Figure 2.   (A) Overlaid UV chromatograms of labeled sweat samples collected after 10-20, 20-30 min and 30-40 min exercise and labeled method blank. (B) Average peak areas of labeled sweat collected from all volunteers at different times during exercise and labeled method blank. Error bars represent standard deviation (N=36 for samples and N=3 for blank). (C) LC-MS base-peak chromatogram from a 20 nmol injection of dansyl labeled sweat. Amino acids are indicated by their single letter code.

Figure 3.   PLS-DA plot of sweat samples from different individuals. Each colour represents a different individual's sweat. There are 18 data points for each individual (3 samples $\times$ 3 days $\times$ duplicate labeling). $R^2X = 0.869$, $R^2Y = 0.961$ and $Q^2 = 0.931$.

Figure 4.   (A) PCA plot and (B) PLS-DA plot of sweat samples from males and females ($R^2X = 0.647$, $R^2Y = 0.973$ and $Q^2 = 0.960$). (C) Volcano plot with highlighted features fitting the criteria of $0.67 \geq$ fold-change (FC) ≥ 1.5 and $p \leq 0.05$.

Figure 5.   (A) PCA plot and (B) PLS-DA plots of sweat samples collected after different exercise times ($R^2X = 0.865$, $R^2Y = 0.978$, $Q^2 = 0.701$).

Figure 6.   Volcano plots of sweat collected after different exercise times: (A) T1 vs. T2, (B) T2 vs. T3 and (C) T1 vs. T3. The fold change represents the average ratio of the later exercise period divided by that of the earlier exercise period. Highlighted features fit the criteria of $0.67 \geq FC \geq 1.5$ and $p \leq 0.05$.
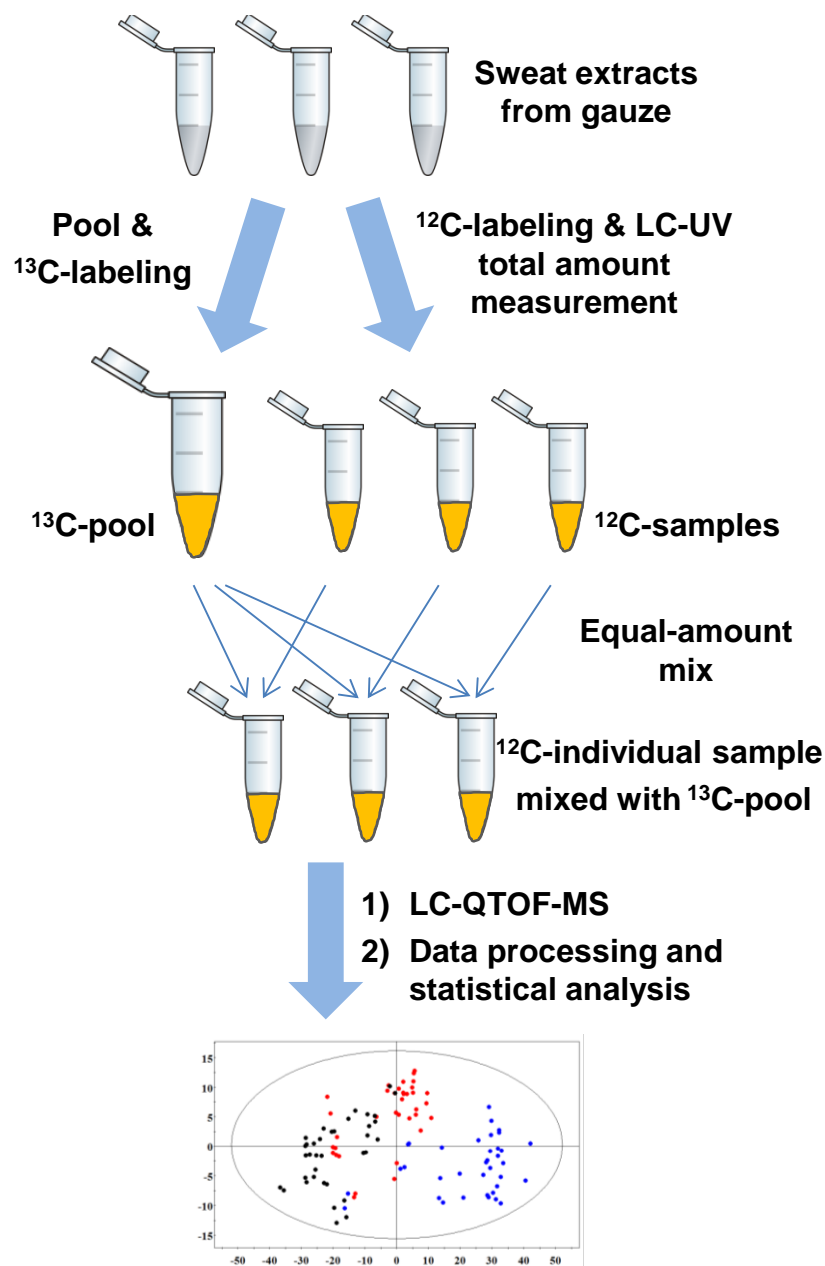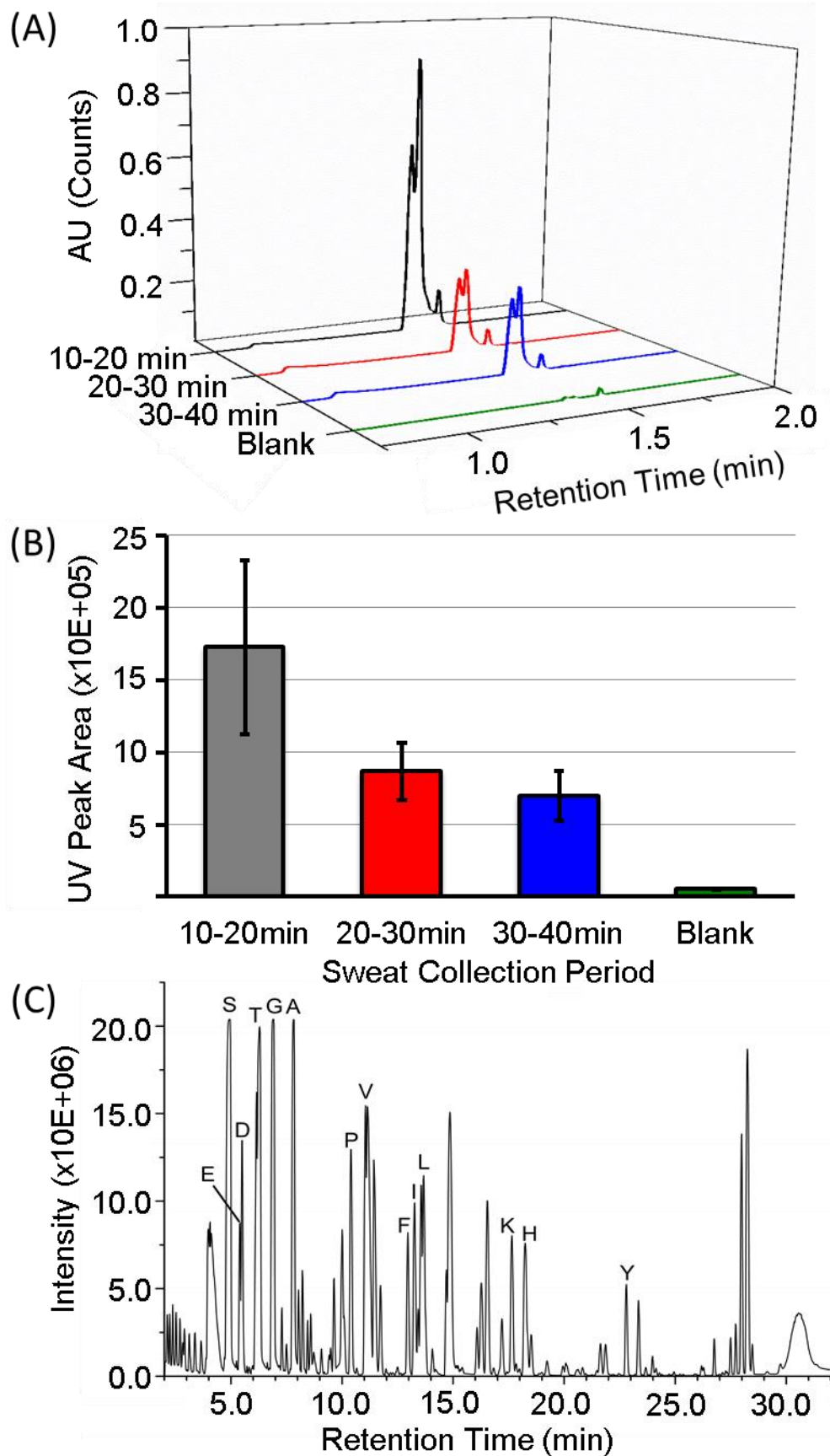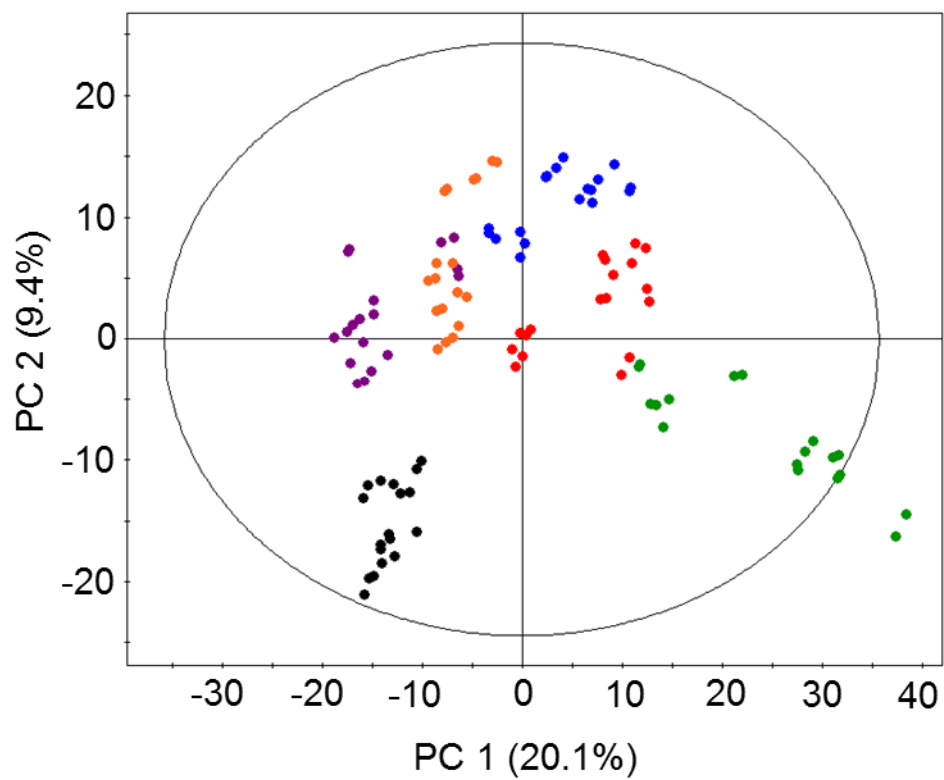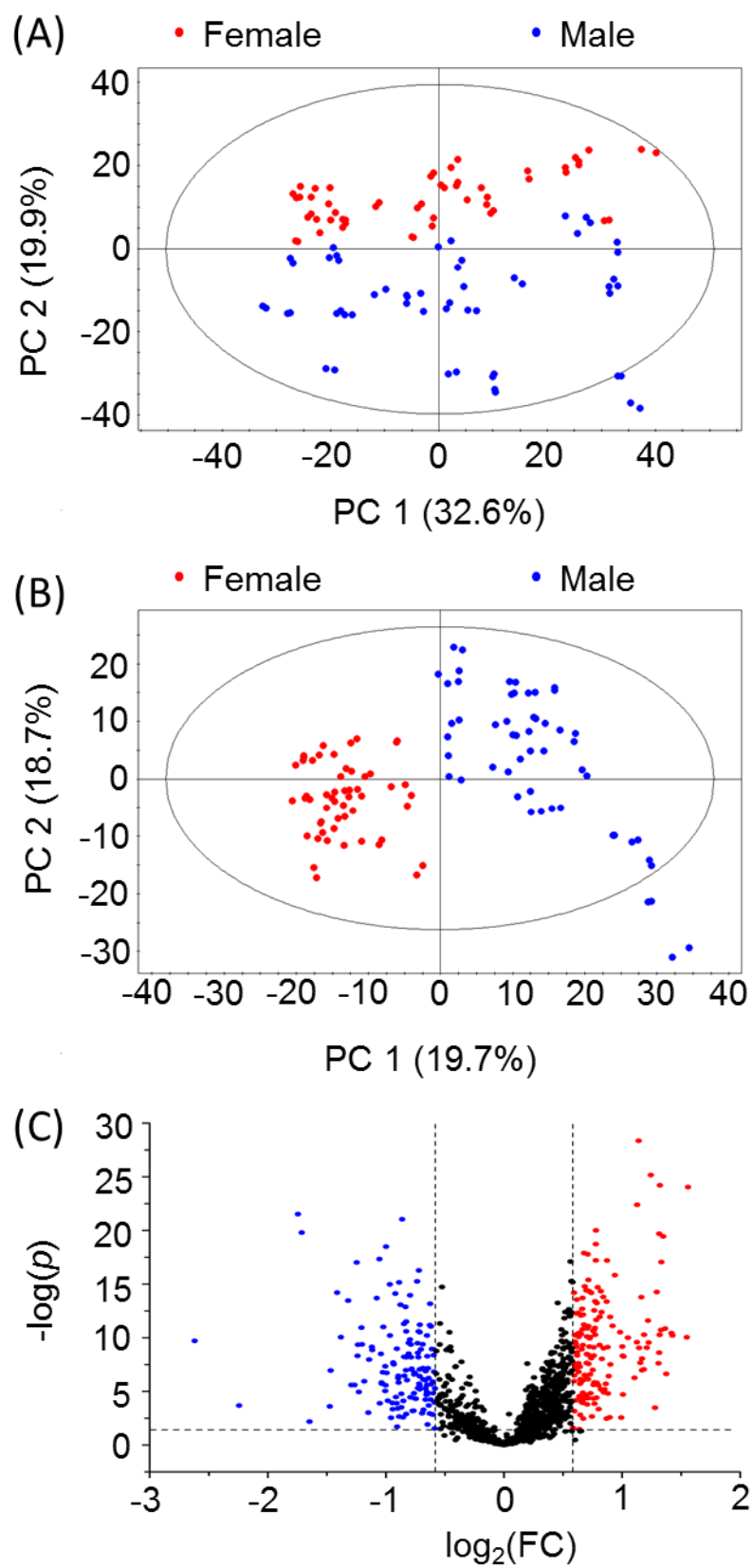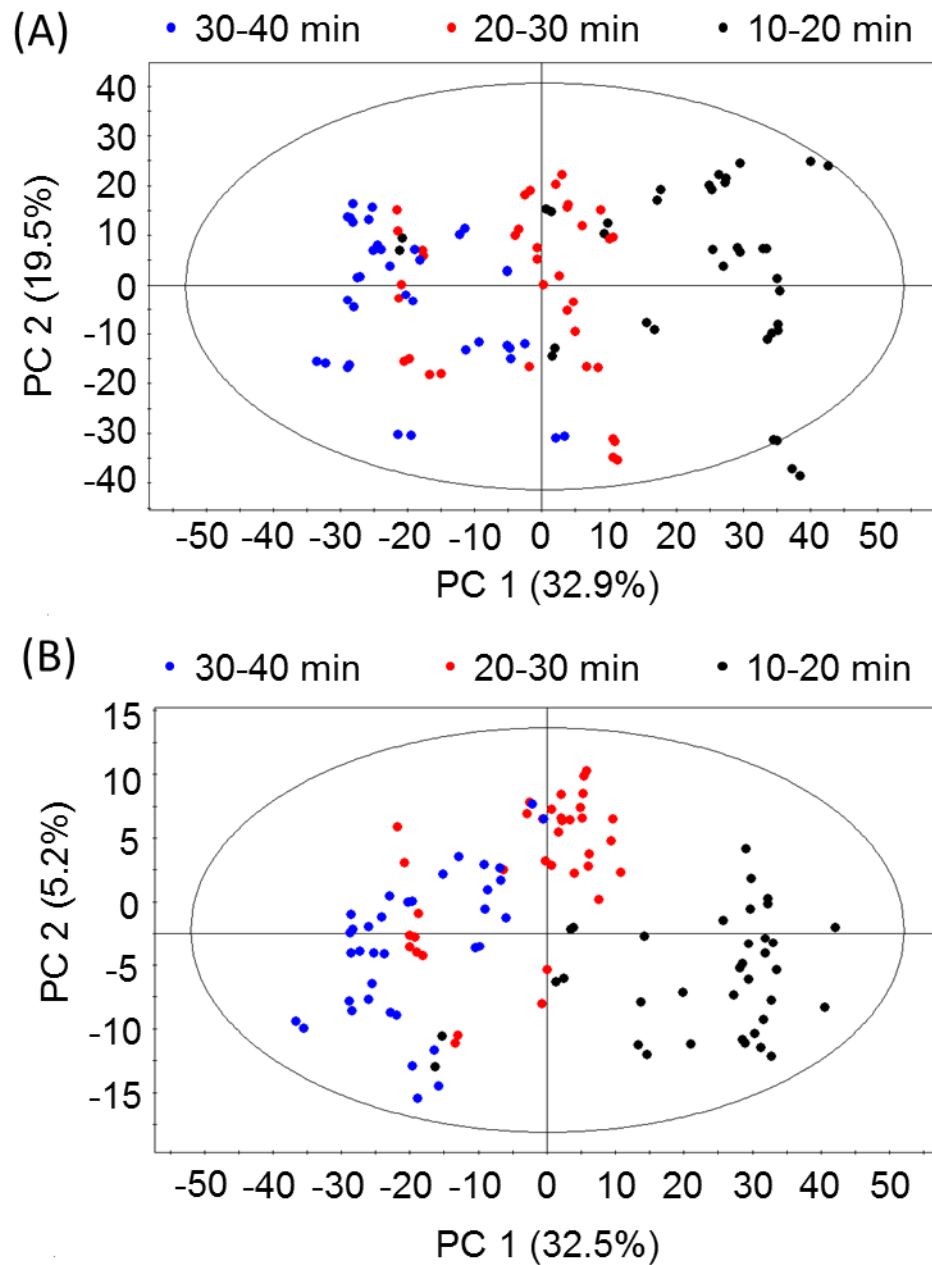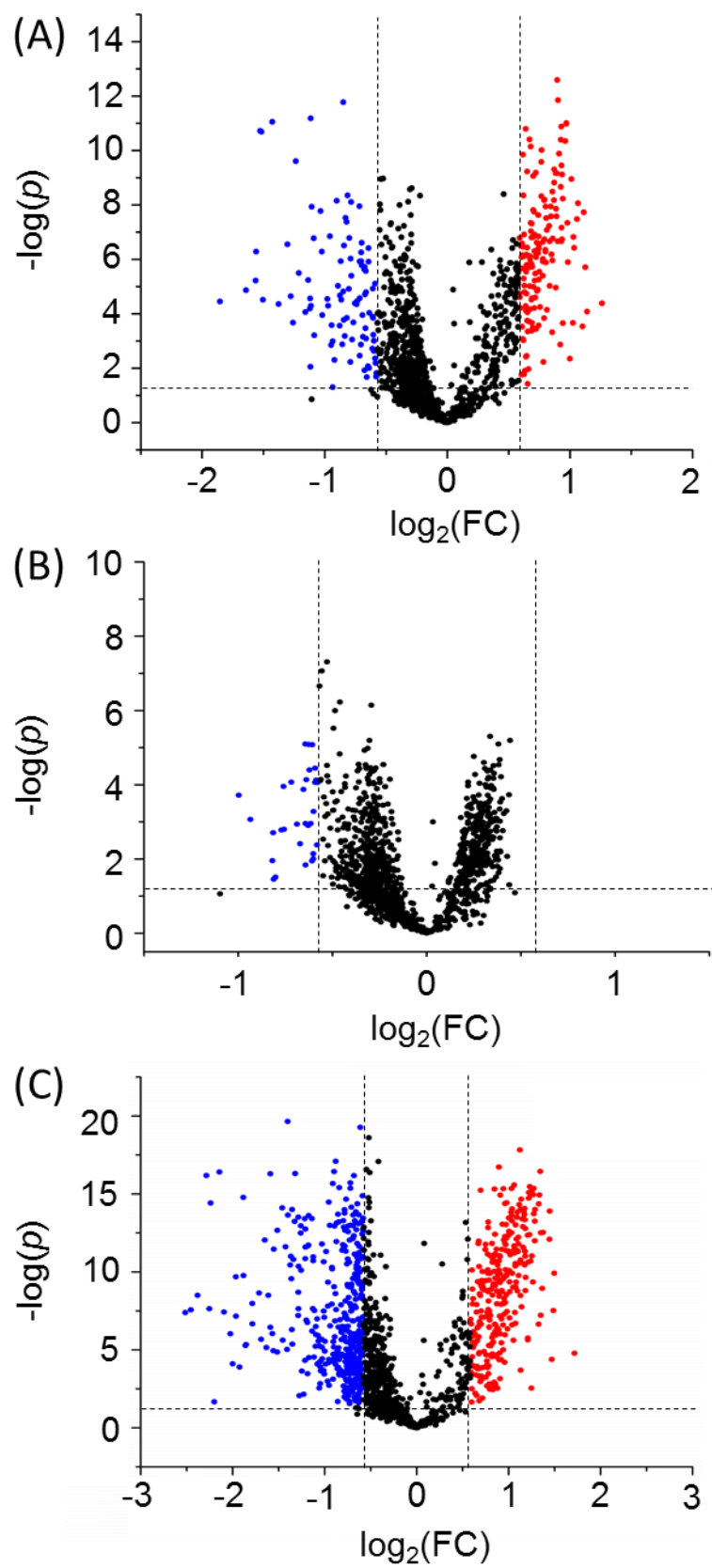
Figure 1

Figure 2

Figure 3

Figure 4

Figure 5

Figure 6

For Table of Contents:

CIL LC-MS for Human Sweat Metabolomics