Received: 9 January 2019 | Revised: 5 May 2019 | Accepted: 22 May 2019

DOI: 10.1002/mgg3.1738

Molecular Genetics & Genomic Medicine Open Access WILEY

ORIGINAL ARTICLE

Individual variability in human urinary metabolites identifies age-related, body mass index-related, and sex-related biomarkers

Tianling Wang[1,2] | **Lei Tang**[1] | **Ruili Lin**[1] | **Dian He**[1,3] | **Yanqing Wu**[1] | **Yang Zhang**[1,4] | **Pingrong Yang**[1,3] | **Junquan He**[1,3]

[1]Materia Medica Development Group, Institute of Medicinal Chemistry, Lanzhou University School of Pharmacy, Lanzhou, China

[2]Dingxi Campus of Gansu, University of Traditional Chinese Medicine, Dingxi, China

[3]Gansu Institute for Drug Control, Lanzhou, China

[4]School of Pharmaceutical Sciences, Chongqing University, Chongqing, China

Correspondence
Dian He and Yang Zhang, Materia Medica Development Group, Institute of Medicinal Chemistry, Lanzhou University School of Pharmacy, Lanzhou 730000, China.
Email: hed@lzu.edu.cn (D. H.); 20162901007@cqu.edu.cn (Y. Z.)

Funding information
Lanzhou Science and Technology Bureau Program Funds, Grant/Award Number: 2017-RC-16 and 2018-1-110; Gansu Science and Technology Fund, Grant/Award Number: 17ZD2FA009

Abstract

Background: Metabolites present in human urine can be influenced by individual physiological parameters (e.g., body mass index [BMI], age, and sex). Observation of altered metabolites concentrations could provide insight into underlying disease pathology, disease prognosis and diagnosis, and facilitate discovery of novel biomarkers.

Methods: Quantitative metabolomics analysis in the urine of 183 healthy individuals was performed based on high-resolution liquid chromatography–mass spectrometry (LC–MS). Coefficients of variation were obtained for 109 urine metabolites of all the 183 human healthy subjects.

Results: Three urine metabolites (such as dehydroepiandrosterone sulfate, acetaminophen glucuronide, and *p*-anisic acid) with $CV_{183} > 0.3$, for which metabolomics studies have been scarce, are considered highly variable here. We identified 30 age-related metabolites, 18 BMI-related metabolites, and 42 sex-related metabolites. Among the identified metabolites, three metabolites were found to be associated with all three physiological parameters (age, BMI, and sex), which included dehydroepiandrosterone sulfate, 3-methylcrotonylglycine and *N*-acetyl-aspartic acid. Pearson's coefficients demonstrated that some age-, BMI-, and sex-related compounds are strongly correlated, suggesting that age, BMI, and sex could affect them concomitantly.

Conclusion: Metabolic differences between distinct physiological statuses were found to be related to several metabolic pathways (such as the caffeine metabolism, the amino acid metabolism, and the carbohydrate metabolism), and these findings may be key for the discovery of new diagnostics and treatments as well as new understandings on the mechanisms of some related diseases.

KEYWORDS
aging markers, BMI markers, CV value, metabolomic, sex markers, urine metabolite

1 | INTRODUCTION

Urine is an easily accessible and well-investigated biological fluid for disease-associated studies (Kostidis et al., 2018; Weiss & Kim, 2011), for example, serial sampling to monitor disease and therapeutic response (Zhang et al., 2012). Human urine metabolites are the downstream products (Wishart, 2016) and end point markers of various biologic processes in the human body (Suhre et al., 2016), which can be a very sensitive measure of an organism's phenotype (Bouatra et al., 2013; Zhang et al., 2015). Observation of the altered urine metabolites concentrations can provide insights into underlying disease pathology (Wild et al., 2013) and disease prognosis and diagnosis (Houten, 2009; Suhre et al., 2016), and facilitate the discovery of novel biomarkers (Perreault et al., 2016). As reported, clinical metabolomics will be the next stage of clinical biochemistry (D'Alessandro et al., 2012). In particular, clinical biochemistry was a research field mainly relied on the biochemical analyses of various biological fluids (e.g., urine, serum, and cerebrospinal fluid; D'Alessandro et al., 2012), for example, Richard Bright's (1789–1858) test. Currently, the introductions of the cutting-edge instrumentation and technology innovation have enabled decades of substantial improvements in the field of standard analytical chemistry in the clinical analysis. For example, the emerging metabolomic technology has been reported to be likely enough add up to the analytical approaches at disposal of clinical analysis.

Metabolomics is a branch of omics science which systematically studies metabolites in organisms, cells, and biofluid (e.g., urine; Chaleckis et al., 2016; Saccenti et al., 2014). Among many emerging metabolomic techniques, the liquid chromatography–mass spectrometry (LC–MS) has advanced advantages of sensitively and simultaneously detecting thousands of metabolites, and has no chemical derivatization requirements and high-throughput capacity (Chen et al., 2007; Dunn et al., 2011; Hildebrandt et al., 2011), which is the main reason that it has been the most widely employed technique in untargeted metabolomic studies of human population (Luan et al., 2015; Mapstone et al., 2014; Torres-Benitez et al., 2017; Wang et al., 2011).

Urinary metabolites were found to be associated with several physiological parameters (e.g., body mass index [BMI], age, or sex; Geifman et al., 2013; Rescigno et al., 2017; Sugimoto et al., 2013; Wu & Gao, 2015). These physiological parameters may influence the filtering of metabolites in the glomeruli and reabsorption in the proximal tubules of the nephron (Wu & Gao, 2015). In addition, physiological parameters could be associated with a higher risk of multiple diseases (e.g., cancer, cardiovascular disease, or neurodegeneration; Johnsson et al., 2018; Joo et al., 2018; Niccoli & Partridge, 2012; Reckelhoff, 2001), which were the key influential factors when considering phenotypic changes in health and disease (Martinez-Selles et al., 2018). For example, a patient's age can affect the course and progression of a disease (Diamond et al., 1989; Geifman & Rubin, 2012; Hasenclever & Diehl, 1998; Joo et al., 2018) or play an important role in determining the correct course of the treatment (Vecht, 1993). Thus, quantitative analysis of a metabolite (or compound) of among individuals offer profound insights into health or disease conditions and the effects of nutrition, drugs, and stress (Chaleckis et al., 2016; Masike et al., 2017). Moreover, comprehensive information about individual variations in metabolites could impact the future of medical science (Patti et al., 2012; Ramautar et al., 2013). However, individual variability and physiological variations of urine metabolite have not been extensively studied, not to mention the potential mechanisms underlying how metabolites alternate between different physiological conditions (Slupsky et al., 2007), which will be the key to the discovery of new diagnosis and treatment and new understandings of the mechanism of disease (Goveia et al., 2016; Wishart, 2016).

Herein, in this study, individual variability in human urine metabolites based on LC–MS untargeted metabolomic study was systematically investigated. The coefficient of variation (CV) is a well-known metric for exploring the metabolites variability among individuals. First, to quantify individual variation, CVs for 109 urine metabolites were calculated and categorized into two distinct groups: low variability and high variability. Second, the statistical significance of metabolites between the distinct studied physiological parameters was calculated by a linear model and moderated *t*-statistics using an empirical Bayes method implemented in the limma R/Bioconductor package. In this study, the adjusted *p* values were estimated using the Benjamini–Hochberg method. Combining the present quantitative data with age, BMI, and sex information, we identified 30 age-related metabolites, 18 BMI-related metabolites, and 42 sex-related metabolites. Among the identified metabolites, three metabolites were found to be associated with all three physiological parameters, which included dehydroepiandrosterone sulfate, 3-methylcrotonylglycine, and *N*-acetyl-aspartic acid. Third, correlation analysis suggested that certain age-related, BMI-related, and sex-related compounds were highly correlated. Moreover, based on these related metabolites, several metabolic pathways (such as caffeine metabolism, amino acid metabolism, and carbohydrate metabolism) were found to be related to age, BMI, or sex.

2 | MATERIALS AND METHODS

2.1 | Subjects in this study

Samples were collected from a previous study of healthy adult urines (Thevenot et al., 2015), which included 183 healthy adults. Samples were obtained with informed consent of the

subjects, which was in accordance with the 1964 Helsinki Declaration. The benchmark dataset was obtained by non-targeted high-resolution liquid chromatography–mass spectrometry (LC–MS) analysis, which was obtained by the "ropls" package. This package is publicly available from the Bioconductor repository. In particular, a total of 109 urine metabolites were annotated for each sample. The annotation of the metabolites was performed by using the HMDB, KEGG, and METLIN public databases, as well as in-house ESI-mass spectra database developed by Thevenot et al. (2015). All detailed descriptions on LC–MS analysis and data acquisition and data preprocessing are provided in the Supplementary Method.

2.2 | Identification of the performance relationship among CVs for each metabolite

The hierarchical clustering (Libbrecht & Noble, 2015) was adopted to differentiate the individual variability among the 109 urine metabolites (Supplementary Method). As previously reported, the CV is a well-known metric for quantifying individual variation of metabolites among healthy subjects. Individually different relative ratios of peak areas were relevant to the obtained CVs. Thus, in this study, the CV was applied for estimating the variability of each urine metabolite. The CV is the ratio of the standard deviation (SD) of metabolite abundance (peak areas) divided by the mean. The high CVs of metabolites could indicate metabolite variation among individuals (Chaleckis et al., 2016). First, the CVs for each urine compound from all the 183 volunteers (CV_{183}) were calculated based on these relative peak areas among subjects using SD mean ratio (Chaleckis et al., 2016). The CV_{183} of a specific metabolite among 109 urine metabolites were used to generate a 109-dimensional vector. Second, hierarchical clustering was adopted to investigate the relationship among the vectors, and therefore among corresponding CV_{183}. To measure the distance between any two vectors, the Manhattan was applied (Kim et al., 2016). The clustering method applied is the Ward's minimum variance, which can reduce the total within-cluster variance to the maximum extent. In this work, Ward's minimum variance module in R package was used (Tippmann, 2015). Finally, a dendrogram visualization of the hierarchical clustering was generated by the version 3 of Interactive Tree Of Life (iTOL) v3 software (Letunic & Bork, 2016).

2.3 | Statistical analysis for selecting changed metabolites between two distinct physiological conditions

To identify the differential age-, BMI-, and sex-related compounds, the linear models for microarray data (limma; McGeachie et al., 2015) and partial least square discriminant analysis (PLS-DA) were applied in this study (Kim et al., 2009; Supplementary Method). The limma is a package for the analysis of gene expression data arising from microarray or RNA-Seq techniques (Ritchie et al., 2015). A core capability is the use of linear models to assess differential expression in the context of multifactor designed experiments (Ritchie et al., 2015). In recent, Manuela et al. also applied a method based on generalized linear model to investigate the relationship between sex- and age-related physiological conditions for human metabolites. Meanwhile, the PLS-DA is a chemometrics technique used for classification purposes either to infer variables that maximize the discrimination between predefined sample groups or to predict class affiliations of unclassified samples based on a calibration set of known class distributions (Bartel et al., 2013).

At this step, first, the abundance values for each metabolite were modeled using a standard fixed effects linear model framework. Second, the *t*-statistic for the physiological conditions effect was then extracted for each metabolite from the selected model and the *p*-values were computed by using one-sided tests. Then, the *p*-values were further adjusted for multiple comparisons by controlling their false discovery rate (FDR; proportion of false positives among the metabolites called significant) at a 5% of threshold. Finally, the metabolites were selected by variable importance in the projection (VIP) values (>1) on the PLS-DA model All procedures were implemented by the "limma" Bioconductor package on the R statistical computing environment (McGeachie et al., 2015).

2.4 | Correlation of urine metabolites identified to be associated with age, BMI, and sex

Pearson's coefficients (Paglia et al., 2015) were applied to examine the correlation among the age-related compounds, the correlation among the BMI-related compounds, and the correlation among the in sex-related compounds. Pearson's coefficients (>0.7; Chaleckis et al., 2016) between two distinct compounds demonstrated that they were strongly correlated, suggesting that aging or BMI or sex conditions could affect them concomitantly.

2.5 | Metabolic pathways revealed by metabolites associated with age, BMI, and sex

To explore the age-, BMI-, and sex-relevant metabolic functions and pathways for understanding the biological meaning of the observed metabolic changes, a metabolite pathway enrichment analysis was conducted using the widely applied MetaboAnalyst 4.0 web tool (Xia et al., 2015). Human

species and hypergeometric test were selected as parameters for the pathway analysis and the threshold of FDR was set at 0.05 for allowing the research discovery (Fanelli et al., 2018).

3 | RESULTS AND DISCUSSION

3.1 | Sample grouping based on age, BMI, and sex

To systematically investigate the variation of urine metabolites, a total of 183 human adults were selected from a previous study, where individuals were sampled by using the Grubbs outlier methodology for being representatives. We performed sample grouping based on the information of age, BMI, and sex (Figure 1). The age was stratified into three age groups according to the commonly used MeSH vocabulary. The teenager is between the ages of 13 and 18, the age of young adults is between 19 and 44, and the middle-aged people is between the ages of 44 to 65 (Geifman et al., 2013). Based on the classification for global database on BMI by the World Health Organization (Allison et al., 2002; Medehouenou et al., 2015), the BMI was categorized into underweight (BMI < 18.5 kg/m^2), normal weight (BMI $18.5 \leq 24.9$ kg/m^2), overweight (BMI $25 \leq 29.9$ kg/m^2), and obese (BMI ≥ 30 kg/m^2; Consultation, 2004; Czwornog & Austin, 2013). Sex was categorized into female and male groups, among which, 100 were males (55%) and 83 were females (45%). Too small sample sizes could reduce the power of the study and increase the probability of error, which can render the study meaningless (Ayeni et al., 2012). Thus, statistical power analysis was performed for estimating sufficient sample sizes to achieve adequate power (Billoir et al., 2015; Blaise et al., 2016). Power values were calculated on all the three datasets at an overall significance level of 1% with Bonferroni's adjustment. Figure S1 provides the statistical power values of the datasets between young adults versus middle-aged (A), normal weight versus overweight people (B), and male versus female individuals (C). For these three datasets, they gave over 90% power to detect differential metabolites at an overall significance level of 0.01 with Bonferroni's adjustment using their corresponding sample sizes. Results suggested that the sample sizes of the datasets were suitable for discriminating metabolites between the young adults and middle-aged groups, between the normal weight and overweight groups, and between the male and female groups.

3.2 | Determination of individual CVs for each urine metabolite

We examined individual metabolite variations among the healthy subjects. The CVs for the entire experimental population of 183 individuals were determined for each compound (CV_{183}) in urine. The unsupervised hierarchical clustering method has been used successfully for identifying the relationship among compounds and features (e.g., metabolites or genes; Caesar et al., 2018). In this study, to identify the relationship of variability among the individuals on the 109 urine metabolites detected, we thus applied the hierarchical clustering to CVs and grouped the 109 urines metabolites into three categories (A, B, and C) based on their CV_{183} values (Figure 2). As illustrated in Figure 2, many compounds with CV_{183} less than 0.07 constituted the least variable subset of urine metabolites. The 43 metabolites highlighted in green in Figure 2 belonged to this group (A group). Forty-seven compounds, highlighted in violet in Figure 2, had CV_{183} values from 0.07 to 0.2 and belonged to the second least variable group. 1,7-Dimethyluric acid, a urinary caffeine metabolite, belonged to this group (B group). The findings indicated that these compounds of the intermediate CV group do not fluctuate on individual diet basis. In addition, the remaining 19 compounds highlighted in blue showed CV_{183} values from

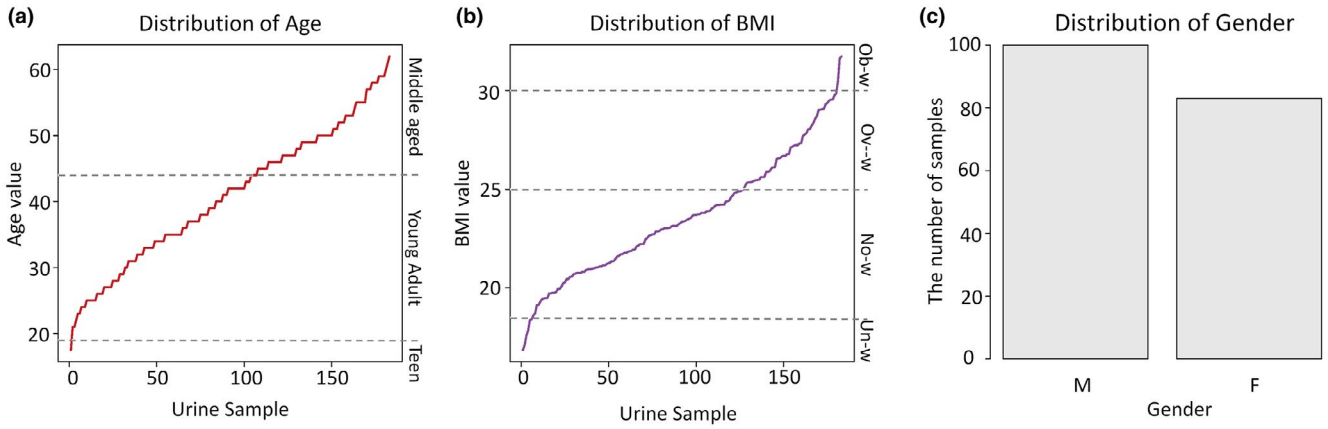

FIGURE 1 Summary of demographic variables of the benchmark dataset. (a) Distribution of age; (b) distribution of BMI; (c) distribution of gender. BMI, body mass index; F, female; M, male; No-w, normal weight; Ov-w, overweight; Ob-w, obese; Teen, Teenagers; Un-w, underweight

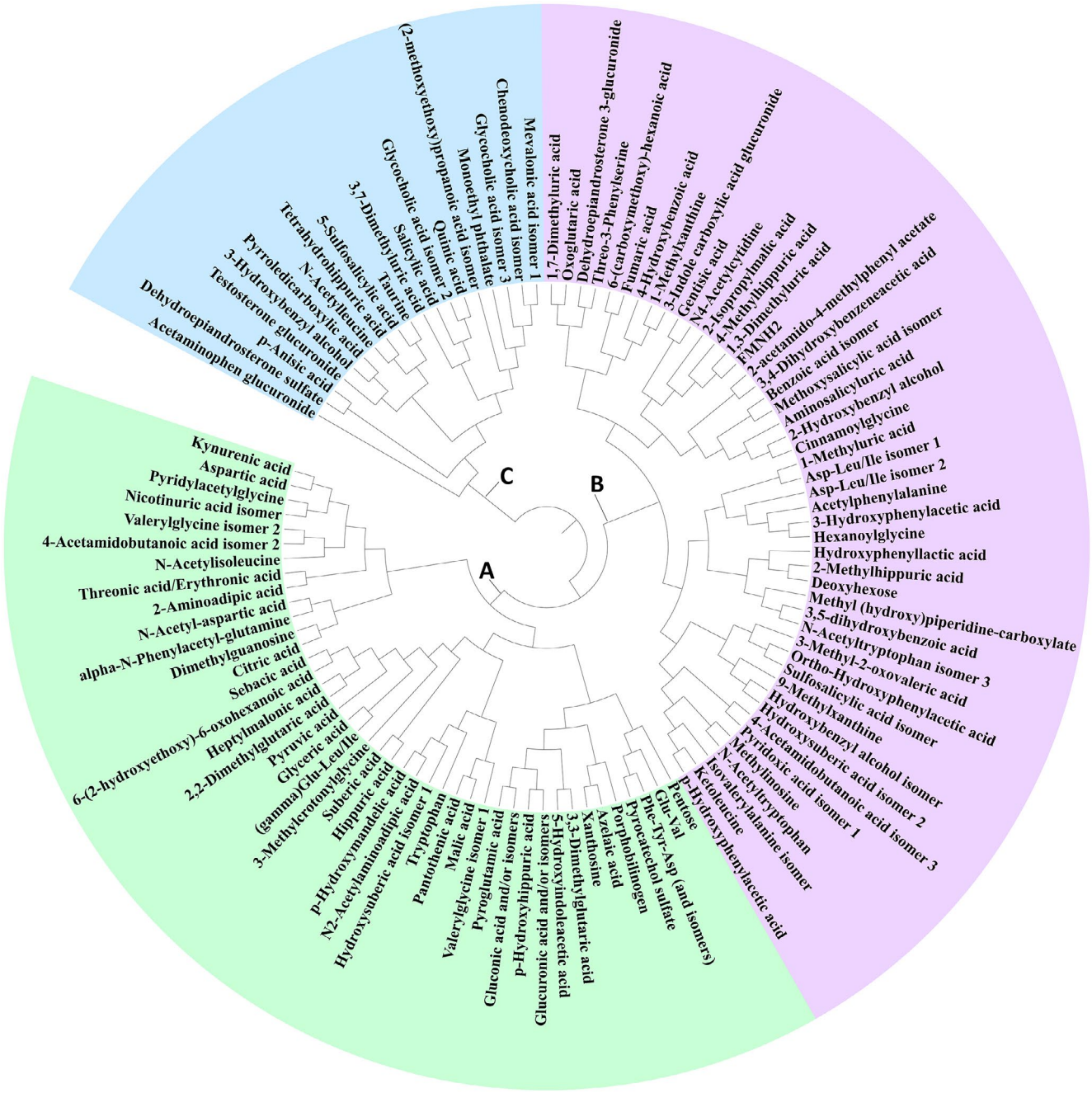

FIGURE 2 Cluster profiling of CV profiles for 109 human urine metabolites. The 109 urine compounds with coefficients of variation (CV_{183}) in three different ranges. The lowest CV_{183} (<0.07) group contains 43 compounds are highlighted in green (a). The middle CV_{183} (0.07–0.3) group contains 47 compounds are highlighted in violet (b). The highest CV_{183} (>0.3) group contains 19 compounds are highlighted in blue (c). The higher CV_{183} indicate high variability groups, the lower CV_{183} indicate low variability groups

0.2 to 1.08. We consider these compounds to be of intermediate to high variability. Among them, 16 compounds had CV_{183} values from 0.2 to 0.3 that could be considered as moderately variable. As an example, testosterone glucuronide, a natural human metabolite of testosterone (Ambrosini et al., 2012) belonged to this group (C group). The compounds with $CV_{183} > 0.3$ were considered highly variable. The three metabolites, dehydroepiandrosterone sulfate, acetaminophen glucuronide and *p*-anisic acid, belonged to the C group; $CV_{183} > 0.3$. To our knowledge, CVs of these three urine metabolites identified had not been previously reported, which could be considered as potential biomarkers for different individuals. On the contrary, compounds with low CVs (A group) may support physiological homeostasis in vivo thereby might be consider as good candidates to represent health biomarkers (Chaleckis et al., 2016).

3.3 | Biological functions analysis for urine metabolites revealed by CV measurements

We classified the 109 detected compounds into 41 categories based on their molecular biological functions. Among the 109 compounds, 71 metabolites could be matched into HMDB database, and many had the information about their biological function. In total, 41 biological functions sets were identified, 12 compounds belonging to the lowest CV_{183} (<0.07) group A were related to 36 biological function sets; 7 compounds belonging to the middle CV_{183} (0.07–0.2) B group were related to 9 biological function sets; and 2 compounds in the C group with the highest CV_{183} (0.2–1.08; e.g., dehydroepiandrosterone sulfate, CV_{183} 0.32; acetaminophen glucuronide, CV_{183} 1.08) were identified to be associated with five biological functions sets: waste products, cell signaling, fuel and energy storage, fuel or energy source, and membrane integrity stability.

Most of the urine compounds (109) could be classified into high-area MS peaks (>5), medium-area MS peaks (4–5), and low-area MS peaks (<4). It is noteworthy that acetaminophen glucuronide (HMDB10316) data were dubious with a CV_{183} of 1.08. A close review of its dataset revealed that seven individuals presented high levels, seven presented middle levels, whereas the remaining of the individuals presented negligible levels of HMDB10316 (Figure S2). HMDB10316 (acetaminophen glucuronide) is a natural human metabolite of acetaminophen generated in the liver by UDP glucuonyltransferase (Wishart et al., 2018), which is involved in the acetaminophen metabolism pathway. As reported in Court's pioneer study, the race, sex, and genetic polymorphism could contribute to variability in acetaminophen metabolism in healthy volunteers (Court et al., 2017). The abundant differences in HMDB10316 across individuals might be explained by the perturbed acetaminophen metabolism pathway (Wishart et al., 2018). Moreover, the very large variability among intensities (HMDB10316) was found to be associated with some metabolic waste products (e.g., the excretion of drugs, toxic substances, or other substances cannot be used as an energy source; Aw & Jones, 1983; Wishart et al., 2018).

3.4 | Evaluation and validation of the "limma" Package approach based on spiked metabolites

A good feature selection method should have the characteristic of selecting differential features that are related to the spiked metabolites (true positives). Thus, in this work, we performed the accuracy evaluation of the "limma" approach by measuring its ability to whether select the true positive spiked-in metabolic features set. First, a benchmark spike-in dataset (MTBLS59) from Franceschi's work (Franceschi

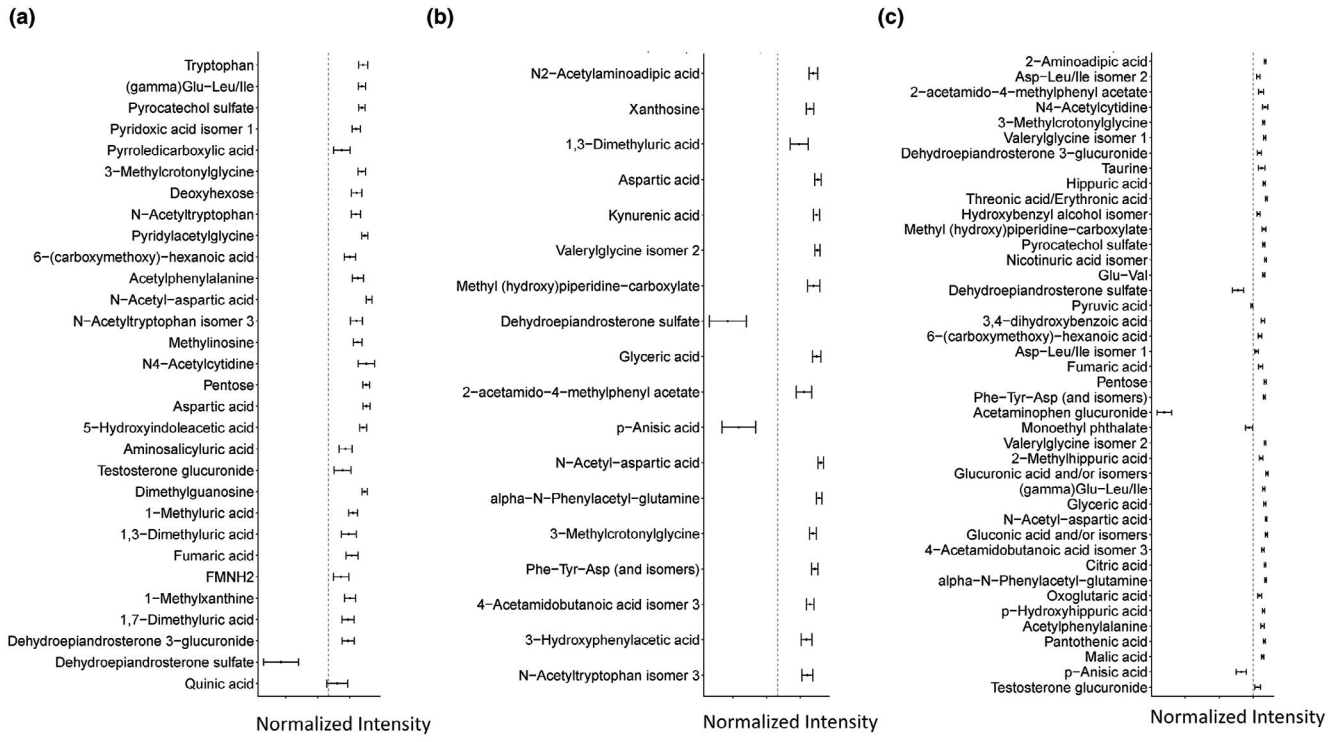

FIGURE 3 Identification of significant differential urine metabolites between the young adults and middle-aged people groups (a), between the normal weight and overweight groups (b), and between the male and female groups (c). Age: young adult (19–44 years of age) and middle-aged people (44–65 years of age); BMI: normal weight (BMI 18.5 ≤ 24.9 kg/m^2), overweight (BMI 25 ≤ 29.9 kg/m^2); Gender: female and male. All differential metabolites were order according to the adjusted *p*-value (adjP) which were calculated using the Benjamini and Hochberg method

et al., 2012) was analyzed (Supplementary Method). Then, a well-suited metric (the number of true positives metabolic features; Christin et al., 2013) was calculated in order to evaluate the accuracy performance of the method. As shown in Table S1, the true positives metabolic features could be identified based on the limma approach, which indirectly reflected the reliability of the strategy applied in this study. Moreover, the PLS-DA method (Kim et al., 2009) adopted for supporting the interpretation of the above results were also further evaluated and validated. The performances on identifying spike-in compounds based on the PLS-DA approach was also consistent with Franceschi's work (Franceschi et al., 2012; Table S1).

3.5 | Discriminating metabolites between young adults and middle-aged people

To identify significantly differential urine metabolites between two distinct age groups, the "limma" approach was used, which could be implemented with the "limma" package (Ritchie et al., 2015). As illustrated in Figure 3a, we found 30 compounds that differed significantly between the two distinct age groups (FDR $\leq$ 0.05). For example, quinic acid (Figure 4a) showed strikingly lower levels in young healthy adults compared with middle-aged healthy people (adjP = 2.42E-07). The levels of two urinary metabolites including the 1,7-dimethyluric acid (adjP = 9.48E-06; Figure 4b) and 1-methylxanthine (adjP = 0.00015; Figure 4c) were clearly less abundant in the young adult people. Similarly, fumaric acid, an organic dicarboxylic acid that played a role in the tricarboxylic cycle (TCA cycle), showed obvious difference between the middle-aged people and the young adults (much less in young adult urine; adjP value of 0.00017; Figure 4d). Meanwhile, the levels of dehydroepiandrosterone sulfate (adjP = 4.52E-06; Figure 4e), dehydroepiandrosterone 3-glucuronide (adjP = 4.52E-06; Figure 4f), FMNH2 (adjP = 0.00016; Figure 4g), and dimethyl-guanosine (adjP = 0.00048; Figure 4h) were clearly more abundant in the urine of the young adults. Moreover, we noted that the discriminating metabolites selected based on the PLS-DA

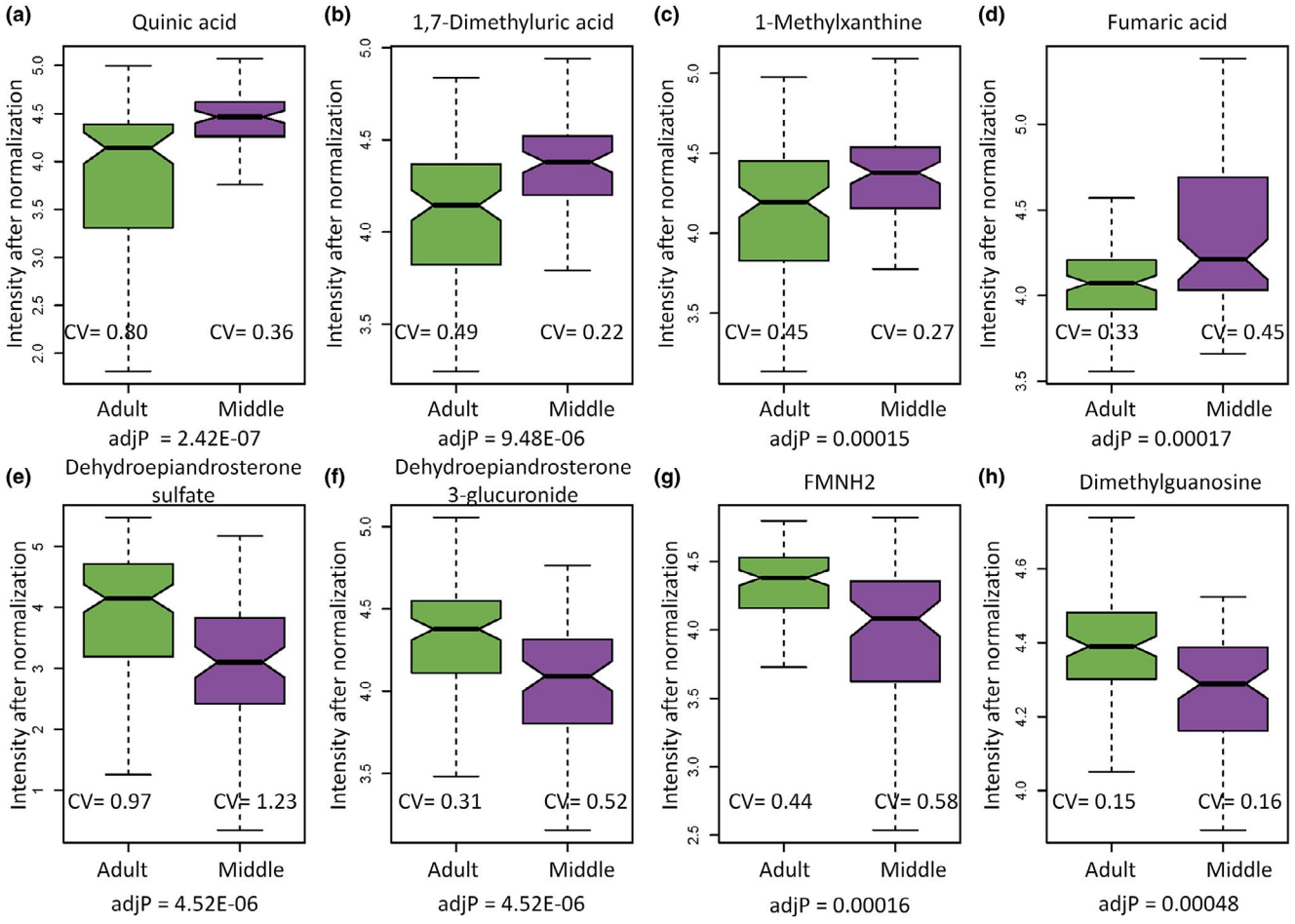

FIGURE 4 Identification of some urine metabolites that differ in abundance between young adult (19–44 years of age) and middle-aged (44–65 years of age) people. Quinic acid (a), 1,7-dimethyluric acid (b), 1-methylxanthine (c), and fumaric acid (d) are higher in middle-aged subjects whereas dehydroepiandrosterone sulfate (e), dehydroepiandrosterone 3-glucuronide (f), FMNH2 (g), and dimethylguanosine (h) are higher in the adult. The CVs indicate individual variability of metabolites. *p* values between the age groups are in the range of 2.42E-07 and .00048

models were present in the ones selected by the "limma" method (Figure S3), which reflected the reliability of the results.

3.6 | Discriminating metabolites between normal weight subjects and overweight subjects

To identify the significant differential urine metabolites associated with BMI, the "limma" approach was used, which could be implemented in the "limma" package (Ritchie et al., 2015). As illustrated in Figure 3b, we found that 18 compounds that differed significantly between the two BMI groups at $p \leq .05$. For example, the methyl-(hydroxy)-piperidine-carboxylate (Figure 5a) showed the lower levels in normal weight subjects compared with overweight ones ($p = .0177$). Kynurenic acid showed impressive difference between the normal weight and overweight individuals (much less in urine normal weight individuals; p value of .0190; Figure 5b) and in agreement with previous studies (Favennec et al., 2015), so high kynurenic acid level was associated with higher BMI ($p < .05$). Similarly, the levels of 1,3-dimethyluric acid ($p = .0280$; Figure 5c) and xanthosine ($p = .0297$; Figure 5d) were clearly less abundant in the normal weight individuals. Meanwhile, the levels of *N*-acetyltryptophan isomer 3 (adjP = 0.0035; Figure 5e), 3-hydroxyphenylacetic acid (adjP = 0.001; Figure 5f), 4-acetamidobutanoic acid isomer 3 (adjP = 0.0165; Figure 5g), and Phe-Tyr-Asp (and isomers; adjP = 0.0231; Figure 5h) were clearly more abundant in the urine of normal weight subjects. The differences in urine metabolites between normal weight and overweight subjects might suggest that certain dysregulated metabolic pathways existed between them.

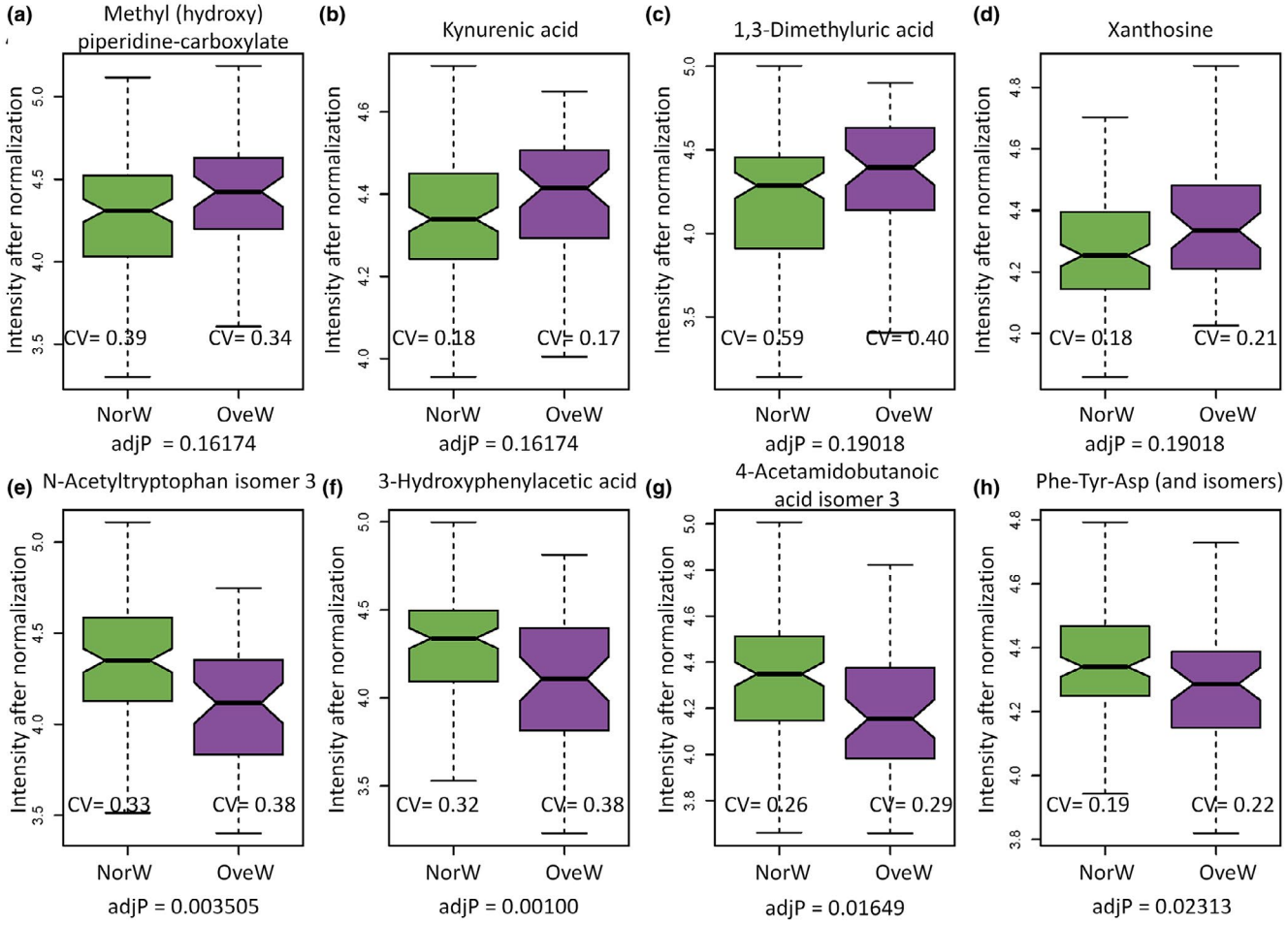

FIGURE 5 Identification of some urine metabolites that differ in abundance between BMI: normal weight (BMI 18.5 ≤ 24.9 kg/m^2) and overweight (BMI 25 ≤ 29.9 kg/m^2) people. Methyl (hydroxy) piperidine-carboxylate (a), kynurenic acid (b), 1,3-dimethyluric acid (c), and xanthosine (d) are higher in overweight subjects, whereas *N*-acetyltryptophan isomer 3 (e), 3-hydroxyphenylacetic acid (f), 4-acetamidobutanoic acid isomer 3 (g), and Phe-Tyr-Asp (and isomers) (h) are higher in the normal weight people. The CVs indicate individual variability of metabolites. p values between the BMI groups are in the range of .003505 and .19018

3.7 | Discriminating metabolites between male and female subjects

To identify the significant differential urine metabolites associated with sex, the "limma" approach was used, which could be implemented with the "limma" package (Ritchie et al., 2015). As illustrated in Figure 3c, we found 42 compounds that differed significantly between the two sex groups (false discovery rate ≤0.05). For example, *p*-anisic acid (Figure 6a) showed strikingly lower levels in healthy male subjects compared with female subjects (adjP = 9.81E-10). Malic acid, a tripeptide analog of glutathione, showed impressive differences between the male and the female groups (much less in the male subjects; adjP value of 4.48E-09; Figure 6b). Similarly, the levels of two oxidant scavengers, pantothenic acid (adjP = 1.31E-07; Figure 6c) and acetylphenylalanine (adjP = 9.23E-06; Figure 6d) were clearly less abundant in the male subjects. Meanwhile, the levels of testosterone glucuronide (adjP = 6.34E-12; Figure 6e), γ-Glu-Leu/Ile (adjP = 0.0012; Figure 6f); Asp-Leu/Ile-isomer-1 (adjP = 0.0059; Figure 6g) and 6-(carboxy-methoxy)-hexanoic acid (adjP = 0.00743; Figure 6h) were clearly more abundant in the urine of male subjects. The differences in urine metabolites between males and females might suggest that certain dysregulated metabolic pathways existed between them.

3.8 | Correlations among age-, BMI-, and sex-related compounds

We found that 13 pairs of 30 age-related compounds that showed relatively strong correlation coefficients (Pearson's *r*; r^2 = 0.70–0.93; Table 1 and Figure S4), 3 pairs of 18 BMI-related compounds showed relatively strong correlation (r^2 = 0.70–0.74; Table 1 and Figure S5) and 12 pairs of 42 sex-related compounds that showed strong correlation (r^2 = 0.70–0.85; Table 1 and Figure S6). Interestingly, among the 22 pairs of age-, BMI-, or sex-related compounds (Table 1), the strongest correlation was found between 1-methyl-uric

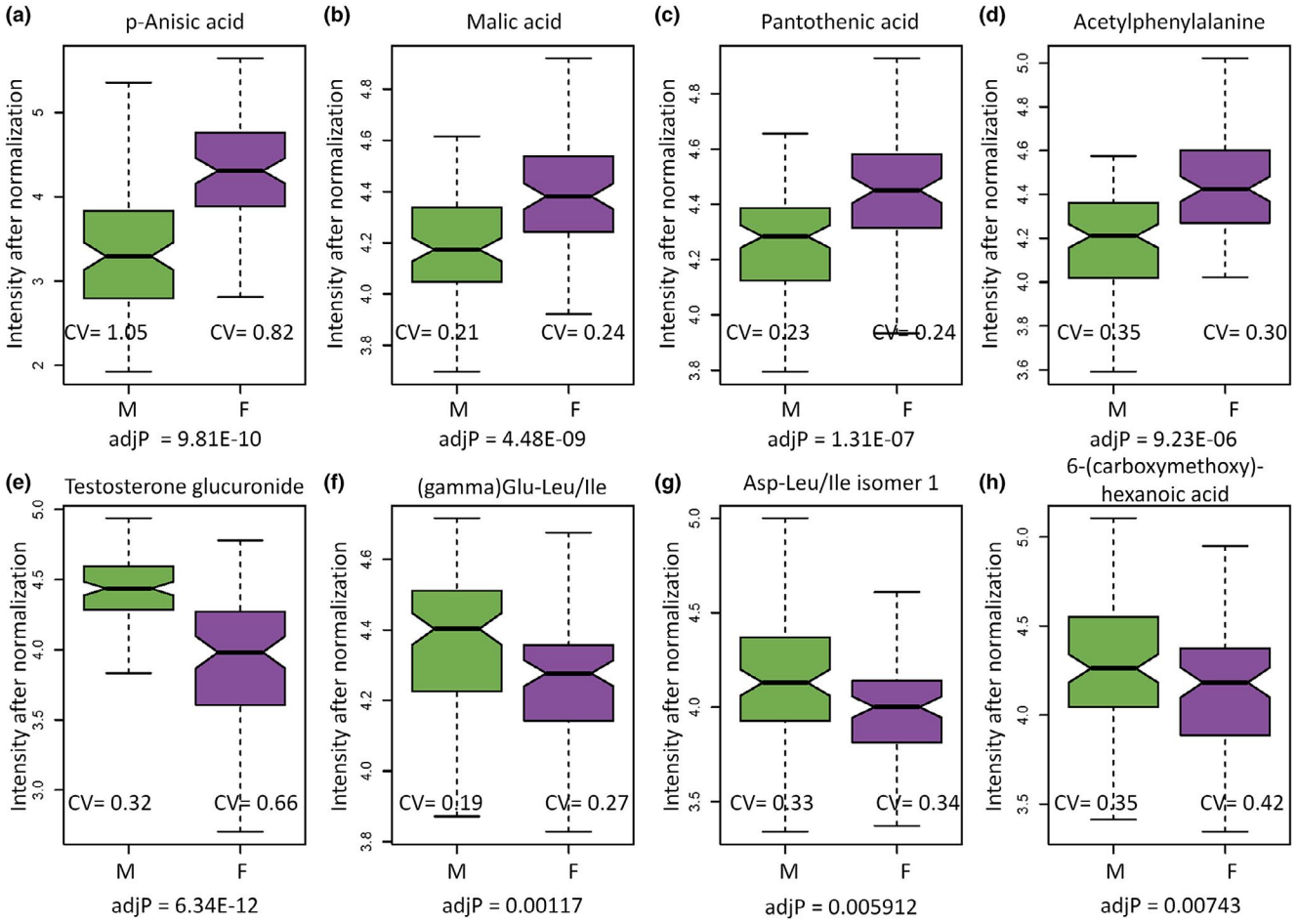

FIGURE 6 Identification of some urine metabolites that differ in abundance between sex: male and female subjects. p-Anisic acid (a), malic acid (b), pantothenic acid (c), and acetylphenylalanine (d) are higher in female subjects, whereas testosterone glucuronide (e), (gamma) Glu-Leu/Ile (f), Asp-Leu/Ile isomer 1 (g), and 6-(carboxymethoxy)-hexanoic acid (h) are higher in the male people. The CVs indicate individual variability of metabolites. *p* values between the gender groups are in the range of 6.34E-12 and .00743

acid and 1-methyl-xanthine (r^2 = 0.927). This may be attributed to their related structures and they are linked in the biochemical pathway of the caffeine metabolism. Potential correlation between dehydro-epiandrosterone-3-glucuronide and testosterone-glucuronide was examined (r^2 = 0.922), and these two compounds are components of the lipid metabolism pathway. Many other correlations between two distinct compounds existed due to similarity chemical taxonomy (e.g., organooxygen compounds or amino acids, peptides, and analogues). In addition, interestingly, among the strong correlation compounds, three pairs of age-related correlations also existed in sex-related pairs: correlation between fumaric acid and pentose (r^2 = 0.796), correlation between pyrocatechol sulfate and pentose (r^2 = 0.724), and correlation between dehydro-epiandrosterone-3-glucuronide and testosterone-glucuronide (r^2 = 0.722). Three pairs of BMI-related correlations also existed in sex-related pairs: correlation between 3-methyl-crotonyl-glycine and valeryl-glycine isomer-2 (r^2 = 0.742), correlation between alpha-*N*-phenyl-acetyl-glutamine and Phe-Tyr-Asp (r^2 = 0.740), and correlation between the 3-methyl-crotonyl-glycine and 4-acetamido-butanoic acid isomer-3 (r^2 = 0.703). The dehydro-epiandrosterone-3-glucuronide and testosterone-glucuronide showed decreased urine levels in the middle-aged people, the latter compound also presented in low levels in female people. Our finding regarding testosterone glucuronide was agreement with a previous study that indicated a sex difference in abundance of testosterone glucuronide, urinary testosterone glucuronide is lower in female subjects (Jones et al., 1977; Perera et al., 1987; Raynaud et al., 1993; Sayo & Hosokawa, 1980). The number of teenagers (1), obese (3), and underweight (6) individuals were too small, and not suited for identifying their regulated metabolites. More urine samples of teenagers, obese, and underweight individuals will be needed for obtaining more general conclusions in the future. Studies aiming elucidate the variability among urine metabolites on the wider age and BMI ranges are of considerable interest nowadays.

TABLE 1 Pairs of age-, BMI-, and gender-related compounds that show relatively high correlation values

Compounds 1	Compounds 2	Correlation	Age-related	BMI-related	Gender-related
1-Methyluric acid	1-Methylxanthine	0.927	+	−	−
1,3-Dimethyluric acid	1-Methyluric acid	0.896	+	−	−
Gluconic acid	Threonic acid	0.851	−	−	+
Hippuric acid	Phe-Tyr-Asp	0.847	−	−	+
1,7-Dimethyluric acid	1-Methylxanthine	0.841	+	−	−
1,3-Dimethyluric acid	1-Methylxanthine	0.832	+	−	−
Gluconic acid	Glucuronic acid	0.828	−	−	+
Glucuronic acid	Threonic acid	0.824	−	−	+
1,3-Dimethyluric acid	1,7-Dimethyluric acid	0.809	+	−	−
Fumaric acid	Pentose	0.796	+	−	+
1,7-Dimethyluric acid	1-Methyluric acid	0.795	+	−	−
Asp-Leu/Ile isomer 1	Asp-Leu/Ile isomer 2	0.762	−	−	+
3-Methylcrotonylglycine	Valerylglycine isomer 1	0.744	−	−	+
3-Methylcrotonylglycine	Valerylglycine isomer 2	0.742	−	+	+
α-*N*-Phenylacetyl-glutamine	Phe-Tyr-Asp (and isomers)	0.740	−	+	+
Pyrocatechol sulfate	Pentose	0.724	+	−	+
Dehydroepiandrosterone 3-glucuronide	Testosterone glucuronide	0.722	+	−	+
1-Methyluric acid	Quinic acid	0.721	+	−	−
Pyridylacetylglycine	Dimethylguanosine	0.711	+	−	−
1-Methylxanthine	Quinic acid	0.704	+	−	−
Deoxyhexose	Fumaric acid	0.704	+	−	−
3-Methylcrotonylglycine	4-Acetamidobutanoic acid isomer 3	0.703	−	+	+

Note: The compounds are selected according to relatively strong correlation coefficients (Pearson's r) (r^2 = 0.70–0.93). These compounds are ordered by the correlation coefficients. Plus (+) indicates that the compounds were correlated in the corresponding condition and minus (−) indicates that the compounds were not correlated in the corresponding condition.

3.9 | Age, BMI, and sex differences in metabolism

To uncover age-, BMI-, and sex-relevant metabolic functions and pathways, we conducted a metabolic pathway enrichment analysis using MetaboAnalyst (Xia et al., 2015). Age-associated metabolites were found to be related to several metabolic pathways (such as caffeine metabolism, and alanine, aspartate, and glutamate metabolism; Table 2). The BMI-associated metabolites were found to be related to the glycine, serine, and threonine metabolism and also the cyanoamino acid metabolism (Table 2). The sex-associated metabolites were found to be related to the carbohydrate metabolism (such as citrate cycle, glyoxylate, dicarboxylate, and butanoate metabolism; Table 2) and the amino acid metabolism (such as phenylalanine metabolism, and alanine, aspartate, and glutamate metabolism; Table 2). Interestingly, age-specific dysregulated metabolic pathways identified (e.g., caffeine metabolism) were operative in age-associated diseases such as Parkinson's disease (Fujimaki et al., 2018; Makrantonaki et al., 2006; Palacios et al., 2010), Alzheimer's disease (Ribeiro & Sebastiao, 2010), and Huntington's disease (Lee & Chern, 2014), which may provide valuable clues to explain the relationship between age and risk of Parkinson's disease. Glycine is a major amino acid in mammals and other animals (Glynn et al., 2015), we also discovered the impact of glycine metabolism in overweight humans, which was consistent with previous published study (Glynn et al., 2015). Moreover, sex-related variability in carbohydrate metabolism was also discovered (Wismann & Willoughby, 2006), which might contribute to 17-β-estradiol mediation, a major determinant of the sex dimorphic response and could affect the organs (Tarnopolsky & Ruby, 2001). In addition, an evidence from previous study suggested the administration of 17-β-estradiol could result in a lower level of glucose and catecholamines in amenorrhoeic women or men (Tarnopolsky & Ruby, 2001). The potential mechanisms behind these sex differences in carbohydrate metabolism need to be further validated.

Moreover, based on a systematic study on urine and serum metabolites by Lau et al. (2018), population-specific variance (age, sex, BMI, ethnicity, dietary, and country of origin) was identified both in urine and serum metabolites, which indicated the ethnicity parameters may be a factor contributing

TABLE 2 Metabolic pathways differed significantly between the two age groups, two BMI groups, and two gender groups

Metabolic pathway	*p*-value	FDR
Age-specific		
Caffeine metabolism	.004	0.237
Alanine, aspartate, and glutamate metabolism	.006	0.237
Nitrogen metabolism	.015	0.302
Nicotinate and nicotinamide metabolism	.019	0.302
Phenylalanine metabolism	.020	0.302
Glycine, serine, and threonine metabolism	.023	0.302
BMI-specific		
Glycine, serine, and threonine metabolism	.008	0.614
Cyanoamino acid metabolism	.045	0.762
Gender-specific		
Citrate cycle (TCA cycle)	4.47E-06	0.0003
Phenylalanine metabolism	.0001	0.005
Glyoxylate and dicarboxylate metabolism	.0002	0.005
Alanine, aspartate, and glutamate metabolism	.0004	0.007
Butanoate metabolism	.002	0.027
Taurine and hypotaurine metabolism	.006	0.086
Pantothenate and CoA biosynthesis	.012	0.129
Pentose phosphate pathway	.016	0.129
Vitamin B6 metabolism	.016	0.129
Lysine biosynthesis	.016	0.129
Nicotinate and nicotinamide metabolism	.029	0.205
Ascorbate and aldarate metabolism	.031	0.205
Glycine, serine, and threonine metabolism	.035	0.213

Note: The pathways are selected as differential according to *p*-value ($p < .05$). All specific pathways are ordered by *p*-value. The false discovery rate (FDR) indicated adjust *p*-value.

to the urine metabolome variance. Therefore, substantial work will be needed to more systematically investigate the ethnicity-specific variance in urine metabolites when the studied subjects are diverse, and the ethnicity information will be fully obtained in the future.

4 | CONCLUSIONS

Human urine provides a rich and important source of information about human metabolites. The coefficient of variation for urine metabolites reflects individual differences in age, BMI and sex conditions. In conclusion, we conducted a comprehensive analysis of urine metabolome physiological variations, and identified 30 age-related metabolites, 18 BMI-related metabolites, and 42 sex-related metabolites. Among them, three metabolites were found to be associated concomitantly with all three physiological parameters (age, BMI, and sex), including dehydroepiandrosterone sulfate, 3-methyl-crotonyl-glycine, and *N*-acetyl-aspartic acid. Correlation analysis suggests that certain age-related compounds, BMI-related compounds, and BMI-related compounds are highly correlated. Individual variability in urine metabolites may also lead to identify candidates for biomarkers of human aging, BMI, sex, or relevant diseases.

ACKNOWLEDGMENTS

The authors thank Thevenot et al. for providing publicly available human urine metabolomic data. This work was funded by the research support of the Lanzhou Science and Technology Bureau Program Funds (2017-RC-16, 2018-1-110) and Gansu Science and Technology Fund grant (17ZD2FA009).

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

Y.Z. and D.H conceived the idea and supervised the work. T.W., L.T., R.L., Y.W., P.Y., and J.H. analyzed data and results. Y.Z. wrote the manuscript. All authors reviewed and approved the manuscript.

ORCID

Yang Zhang https://orcid.org/0000-0002-4909-4939

REFERENCES

Allison, D. B., Zhu, S. K., Plankey, M., Faith, M. S., & Heo, M. (2002). Differential associations of body mass index and adiposity with all-cause mortality among men in the first and second National Health and Nutrition Examination Surveys (NHANES I and NHANES II) follow-up studies. *International Journal of Obesity and Related Metabolic Disorders*, *26*(3), 410–416. https://doi.org/10.1038/sj.ijo.0801925

Ambrosini, S., Shinde, S., De Lorenzi, E., & Sellergren, B. (2012). Glucuronide directed molecularly imprinted solid-phase extraction: Isolation of testosterone glucuronide from its parent drug in urine. *Analyst*, *137*(1), 249–254. https://doi.org/10.1039/c1an15606c

Aw, T. Y., & Jones, D. P. (1983). Intracellular inhibition of UDP-glucose dehydrogenase during ethanol oxidation. *Chemico-Biological Interactions*, *43*(3), 283–288.

Ayeni, O., Dickson, L., Ignacy, T. A., & Thoma, A. (2012). A systematic review of power and sample size reporting in randomized controlled trials within plastic surgery. *Plastic and Reconstructive Surgery*, *130*(1), 78e–86e. https://doi.org/10.1097/PRS.0b013e318254b1d1

Bartel, J., Krumsiek, J., & Theis, F. J. (2013). Statistical methods for the analysis of high-throughput metabolomics data. *Computational and Structural Biotechnology Journal*, *4*, e201301009. https://doi.org/10.5936/csbj.201301009

Billoir, E., Navratil, V., & Blaise, B. J. (2015). Sample size calculation in metabolic phenotyping studies. *Briefings in Bioinformatics*, *16*(5), 813–819. https://doi.org/10.1093/bib/bbu052

Blaise, B. J., Correia, G., Tin, A., Young, J. H., Vergnaud, A.-C., Lewis, M., Pearce, J. T. M., Elliott, P., Nicholson, J. K., Holmes, E., & Ebbels, T. M. D. (2016). Power analysis and sample size determination in metabolic phenotyping. *Analytical Chemistry*, *88*(10), 5179–5188. https://doi.org/10.1021/acs.analchem.6b00188

Bouatra, S., Aziat, F., Mandal, R., Guo, A. C., Wilson, M. R., Knox, C., Bjorndahl, T. C., Krishnamurthy, R., Saleem, F., Liu, P., Dame, Z. T., Poelzer, J., Huynh, J., Yallou, F. S., Psychogios, N., Dong, E., Bogumil, R., Roehring, C., & Wishart, D. S. (2013). The human urine metabolome. *PLoS One*, *8*(9), e73076. https://doi.org/10.1371/journal.pone.0073076

Caesar, L. K., Kvalheim, O. M., & Cech, N. B. (2018). Hierarchical cluster analysis of technical replicates to identify interferents in untargeted mass spectrometry metabolomics. *Analytica Chimica Acta*, *1021*, 69–77. https://doi.org/10.1016/j.aca.2018.03.013

Chaleckis, R., Murakami, I., Takada, J., Kondoh, H., & Yanagida, M. (2016). Individual variability in human blood metabolites identifies age-related differences. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(16), 4252–4259. https://doi.org/10.1073/pnas.1603023113

Chen, C., Gonzalez, F. J., & Idle, J. R. (2007). LC-MS-based metabolomics in drug metabolism. *Drug Metabolism Reviews*, *39*(2–3), 581–597. https://doi.org/10.1080/03602530701497804

Christin, C., Hoefsloot, H. C., Smilde, A. K., Hoekman, B., Suits, F., Bischoff, R., & Horvatovich, P. (2013). A critical assessment of feature selection methods for biomarker discovery in clinical proteomics. *Molecular & Cellular Proteomics*, *12*(1), 263–276. https://doi.org/10.1074/mcp.M112.02256

Consultation, W. H. O. E. (2004). Appropriate body-mass index for Asian populations and its implications for policy and intervention strategies. *Lancet*, *363*(9403), 157–163. https://doi.org/10.1016/S0140-6736(03)15268-3

Court, M. H., Zhu, Z., Masse, G., Duan, S. X., James, L. P., Harmatz, J. S., & Greenblatt, D. J. (2017). Race, gender, and genetic polymorphism contribute to variability in acetaminophen pharmacokinetics, metabolism, and protein-adduct concentrations in healthy African-American and European-American volunteers. *Journal of Pharmacology and Experimental Therapeutics*, *362*(3), 431–440. https://doi.org/10.1124/jpet.117.242107

Czwornog, J., & Austin, G. L. (2013). Body mass index, age, and gender affect prep quality, sedation use, and procedure time during screening colonoscopy. *Digestive Diseases and Sciences*, *58*(11), 3127–3133. https://doi.org/10.1007/s10620-013-2746-2

D'Alessandro, A., Giardina, B., Gevi, F., Timperio, A. M., & Zolla, L. (2012). Clinical metabolomics: The next stage of clinical biochemistry. *Blood Transfusion*, *10*(Suppl 2), s19–s24. https://doi.org/10.2450/2012.005S

Diamond, S. G., Markham, C. H., Hoehn, M. M., McDowell, F. H., & Muenter, M. D. (1989). Effect of age at onset on progression and mortality in Parkinson's disease. *Neurology*, *39*(9), 1187–1190.

Dunn, W. B., Broadhurst, D., Begley, P., Zelena, E., Francis-McIntyre, S., Anderson, N., Brown, M., Knowles, J. D., Halsall, A., Haselden, J. N., Nicholls, A. W., Wilson, I. D., Kell, D. B., & Goodacre, R. (2011). Procedures for large-scale metabolic profiling of serum and plasma using gas chromatography and liquid chromatography coupled to mass spectrometry. *Nature Protocols*, *6*(7), 1060–1083. https://doi.org/10.1038/nprot.2011.335

Fanelli, G., Gevi, F., Belardo, A., & Zolla, L. (2018). Metabolic patterns in insulin-sensitive male hypogonadism. *Cell Death & Disease*, *9*(6), 653. https://doi.org/10.1038/s41419-018-0588-8

Favennec, M., Hennart, B., Caiazzo, R., Leloire, A., Yengo, L., Verbanck, M., Arredouani, A., Marre, M., Pigeyre, M., Bessede, A., Guillemin, G. J., Chinetti, G., Staels, B., Pattou, F., Balkau, B., Allorge, D., Froguel, P., & Poulain-Godefroy, O. (2015). The kynurenine pathway is activated in human obesity and shifted toward kynurenine monooxygenase activation. *Obesity*, *23*(10), 2066–2074. https://doi.org/10.1002/oby.21199

Franceschi, P., Masuero, D., Vrhovsek, U., Mattivi, F., & Wehrens, R. (2012). A benchmark spike-in data set for biomarker identification in metabolomics. *Journal of Chemometrics*, *26*(1), 16–24.

Fujimaki, M., Saiki, S., Li, Y., Kaga, N., Taka, H., Hatano, T., Ishikawa, K.-I., Oji, Y., Mori, A., Okuzumi, A., Koinuma, T., Ueno, S.-I., Imamichi, Y., Ueno, T., Miura, Y., Funayama, M., & Hattori, N. (2018). Serum caffeine and metabolites are reliable biomarkers of early Parkinson disease. *Neurology*, *90*(5), e404–e411. https://doi.org/10.1212/WNL.0000000000004888

Geifman, N., Cohen, R., & Rubin, E. (2013). Redefining meaningful age groups in the context of disease. *Age*, *35*(6), 2357–2366. https://doi.org/10.1007/s11357-013-9510-6

Geifman, N., & Rubin, E. (2012). The age-phenome database. *SpringerPlus*, *1*(1), 1–8. https://doi.org/10.1186/2193-1801-1-4

Glynn, E. L., Piner, L. W., Huffman, K. M., Slentz, C. A., Elliot-Penry, L., AbouAssi, H., White, P. J., Bain, J. R., Muehlbauer, M. J., Ilkayeva, O. R., Stevens, R. D., Porter Starr, K. N., Bales, C. W., Volpi, E., Brosnan, M. J., Trimmer, J. K., Rolph, T. P., Newgard, C. B., & Kraus, W. E. (2015). Impact of combined resistance and aerobic exercise training on branched-chain amino acid turnover, glycine metabolism and insulin sensitivity in overweight humans. *Diabetologia*, *58*(10), 2324–2335. https://doi.org/10.1007/s00125-015-3705-6

Goveia, J., Pircher, A., Conradi, L.-C., Kalucka, J., Lagani, V., Dewerchin, M., Eelen, G., DeBerardinis, R. J., Wilson, I. D., & Carmeliet, P. (2016). Meta-analysis of clinical metabolic profiling studies in cancer: Challenges and opportunities. *EMBO Molecular Medicine*, *8*(10), 1134–1142. https://doi.org/10.15252/emmm.201606798

Hasenclever, D., Diehl, V., Armitage, J. O., Assouline, D., Björkholm, M., Brusamolino, E., Canellos, G. P., Carde, P., Crowther, D., Cunningham, D., Eghbali, H., Ferm, C., Fisher, R. I., Glick, J. H., Glimelius, B., Gobbi, P. G., Holte, H., Horning, S. J., Lister, T. A., … Hudson, G. V. (1998). A prognostic score for advanced Hodgkin's disease. International prognostic factors project on advanced Hodgkin's disease. *New England Journal of Medicine*, *339*(21), 1506–1514. https://doi.org/10.1056/NEJM199811193392104

Hildebrandt, C., Wolf, S., & Neumann, S. (2011). Database supported candidate search for metabolite identification. *Journal of Integrative Bioinformatics*, *8*(2), 157. https://doi.org/10.2390/biecoll-jib-2011-157

Houten, S. M. (2009). Metabolomics: Unraveling the chemical individuality of common human diseases. *Annals of Medicine*, *41*(6), 402–407. https://doi.org/10.1080/07853890902729794

Johnsson, I. W., Naessen, T., Ahlsson, F., & Gustafsson, J. (2018). High birth weight was associated with increased radial artery intima thickness but not with other investigated cardiovascular risk factors in adulthood. *Acta Paediatrica*, *107*(12), 2152–2157. https://doi.org/10.1111/apa.14414

Jones, M. K., Ramsay, I. D., & Collins, W. P. (1977). Concentration of testosterone glucuronide in urine from women with breast tumours. *British Journal of Cancer*, *35*(6), 885–887.

Joo, S. H., Yun, S. H., Kang, D. W., Hahn, C. T., Lim, H. K., & Lee, C. U. (2018). Body mass index in mild cognitive impairment according to age, sex, cognitive intervention, and hypertension and risk of progression to Alzheimer's disease. *Frontiers in Psychiatry*, *9*, 142. https://doi.org/10.3389/fpsyt.2018.00142

Kim, J., Mouw, K. W., Polak, P., Braunstein, L. Z., Kamburov, A., Tiao, G., Kwiatkowski, D. J., Rosenberg, J. E., Van Allen, E. M., D'Andrea, A. D., & Getz, G. (2016). Somatic ERCC2 mutations are associated with a distinct genomic signature in urothelial tumors. *Nature Genetics*, *48*(6), 600–606. https://doi.org/10.1038/ng.3557

Kim, K., Aronov, P., Zakharkin, S. O., Anderson, D., Perroud, B., Thompson, I. M., & Weiss, R. H. (2009). Urine metabolomics analysis for kidney cancer detection and biomarker discovery. *Molecular & Cellular Proteomics*, *8*(3), 558–570. https://doi.org/10.1074/mcp.M800165-MCP200

Kostidis, S., Bank, J. R., Soonawala, D., Nevedomskaya, E., van Kooten, C., Mayboroda, O. A., & de Fijter, J. W. (2018). Urinary metabolites predict prolonged duration of delayed graft function in DCD kidney transplant recipients. *American Journal of Transplantation*, *19*(1), 110–122. https://doi.org/10.1111/ajt.14941

Lau, C.-H., Siskos, A. P., Maitre, L., Robinson, O., Athersuch, T. J., Want, E. J., Urquiza, J., Casas, M., Vafeiadi, M., Roumeliotaki, T., McEachan, R. R. C., Azad, R., Haug, L. S., Meltzer, H. M., Andrusaityte, S., Petraviciene, I., Grazuleviciene, R., Thomsen, C., Wright, J., … Coen, M. (2018). Determinants of the urinary and serum metabolome in children from six European populations. *BMC Medicine*, *16*(1), 202. https://doi.org/10.1186/s12916-018-1190-8

Lee, C. F., & Chern, Y. (2014). Adenosine receptors and Huntington's disease. *International Review of Neurobiology*, *119*, 195–232. https://doi.org/10.1016/B978-0-12-801022-8.00010-6

Letunic, I., & Bork, P. (2016). Interactive tree of life (iTOL) v3: An online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Research*, *44*(W1), W242–W245. https://doi.org/10.1093/nar/gkw290

Libbrecht, M. W., & Noble, W. S. (2015). Machine learning applications in genetics and genomics. *Nature Reviews Genetics*, *16*(6), 321–332. https://doi.org/10.1038/nrg3920

Luan, H., Liu, L.-F., Meng, N., Tang, Z., Chua, K.-K., Chen, L.-L., Song, J.-X., Mok, V. C. T., Xie, L.-X., Li, M., & Cai, Z. (2015). LC-MS-based urinary metabolite signatures in idiopathic Parkinson's disease. *Journal of Proteome Research*, *14*(1), 467–478. https://doi.org/10.1021/pr500807t

Makrantonaki, E., Adjaye, J., Herwig, R., Brink, T. C., Groth, D., Hultschig, C., Lehrach, H., & Zouboulis, C. C. (2006). Age-specific hormonal decline is accompanied by transcriptional changes in human sebocytes in vitro. *Aging Cell*, *5*(4), 331–344. https://doi.org/10.1111/j.1474-9726.2006.00223.x

Mapstone, M., Cheema, A. K., Fiandaca, M. S., Zhong, X., Mhyre, T. R., MacArthur, L. H., Hall, W. J., Fisher, S. G., Peterson, D. R., Haley, J. M., Nazar, M. D., Rich, S. A., Berlau, D. J., Peltz, C. B., Tan, M. T., Kawas, C. H., & Federoff, H. J. (2014). Plasma phospholipids identify antecedent memory impairment in older adults. *Nature Medicine*, *20*(4), 415–418. https://doi.org/10.1038/nm.3466

Martínez-Sellés, M., Díez-Villanueva, P., Álvarez Garcia, J., Ferrero-Gregori, A., Vives-Borrás, M., Worner, F., Bardají, A., Delgado, J. F., Vázquez, R., González-Juanatey, J. R., Fernández-Aviles, F., & Cinca, J. (2018). Influence of sex and pregnancy on survival in patients admitted with heart failure: Data from a prospective multicenter registry. *Clinical Cardiology*, *41*(7), 924–930. https://doi.org/10.1002/clc.22979

Masike, K., Khoza, B. S., Steenkamp, P. A., Smit, E., Dubery, I. A., & Madala, N. E. (2017). A Metabolomics-guided exploration of the phytochemical constituents of vernonia fastigiata with the aid of pressurized hot water extraction and liquid chromatography-mass spectrometry. *Molecules*, *22*(8), 1200. https://doi.org/10.3390/molecules22081200

McGeachie, M. J., Dahlin, A., Qiu, W., Croteau-Chonka, D. C., Savage, J., Wu, A. C., Wan, E. S., Sordillo, J. E., Al-Garawi, A., Martinez, F. D., Strunk, R. C., Lemanske, R. F., Liu, A. H., Raby, B. A., Weiss, S., Clish, C. B., & Lasky-Su, J. A. (2015). The metabolomics of asthma control: A promising link between genetics and disease. *Immunity, Inflammation and Disease*, *3*(3), 224–238. https://doi.org/10.1002/iid3.61

Medehouenou, T. C., Ayotte, P., St-Jean, A., Meziou, S., Roy, C., Muckle, G., & Lucas, M. (2015). Overweight and obesity prevalence among school-aged nunavik inuit children according to three body mass index classification systems. *Journal of Adolescent Health*, *57*(1), 31–36. https://doi.org/10.1016/j.jadohealth.2015.03.022

Niccoli, T., & Partridge, L. (2012). Ageing as a risk factor for disease. *Current Biology*, *22*(17), R741–R752. https://doi.org/10.1016/j.cub.2012.07.024

Paglia, G., Sigurjonsson, O. E., Rolfsson, O., Hansen, M. B., Brynjolfsson, S., Gudmundsson, S., & Palsson, B. O. (2015). Metabolomic analysis of platelets during storage: A comparison between apheresis- and buffy coat-derived platelet concentrates. *Transfusion*, *55*(2), 301–313. https://doi.org/10.1111/trf.12834.

Palacios, N., Weisskopf, M., Simon, K., Gao, X., Schwarzschild, M., & Ascherio, A. (2010). Polymorphisms of caffeine metabolism and estrogen receptor genes and risk of Parkinson's disease in men and women. *Parkinsonism & Related Disorders*, *16*(6), 370–375. https://doi.org/10.1016/j.parkreldis.2010.02.012

Patti, G. J., Yanes, O., & Siuzdak, G. (2012). Innovation: Metabolomics: The apogee of the omics trilogy. *Nature Reviews Molecular Cell Biology*, *13*(4), 263–269. https://doi.org/10.1038/nrm3314

Perera, D. M., McGarrigle, H. H., Lawrence, D. M., & Lucas, M. (1987). Amniotic fluid testosterone and testosterone glucuronide levels in the determination of foetal sex. *Journal of Steroid Biochemistry*, *26*(2), 273–277.

Perreault, L., Starling, A. P., Glueck, D., Brozinick, J. T., Sanders, P., Siddall, P., Kuo, M. S., Dabelea, D., & Bergman, B. C. (2016). Biomarkers of ectopic fat deposition: The next frontier in serum lipidomics. *Journal of Clinical Endocrinology and Metabolism*, *101*(1), 176–182. https://doi.org/10.1210/jc.2015-3213

Ramautar, R., Berger, R., van der Greef, J., & Hankemeier, T. (2013). Human metabolomics: Strategies to understand biology. *Current Opinion in Chemical Biology*, *17*(5), 841–846. https://doi.org/10.1016/j.cbpa.2013.06.015

Raynaud, E., Audran, M., Pages, J. C., Brun, J. F., Fedou, C., Chanal, J. L., & Orsetti, A. (1993). Study of urinary excretion of testosterone and epitestosterone glucuronides in children and adolescents. *Pathologie Biologie*, *41*(2), 159–163.

Reckelhoff, J. F. (2001). Gender differences in the regulation of blood pressure. *Hypertension*, *37*(5), 1199–1208.

Rescigno, T., Micolucci, L., Tecce, M. F., & Capasso, A. (2017). Bioactive nutrients and nutrigenomics in age-related diseases. *Molecules*, *22*(1), 105. https://doi.org/10.3390/molecules22010105

Ribeiro, J. A., & Sebastiao, A. M. (2010). Caffeine and adenosine. *Journal of Alzheimer's Disease*, *20*(Suppl 1), S3–S15. https://doi.org/10.3233/JAD-2010-1379

Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., & Smyth, G. K. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Research*, *43*(7), e47. https://doi.org/10.1093/nar/gkv007

Saccenti, E., Hoefsloot, H. C. J., Smilde, A. K., Westerhuis, J. A., & Hendriks, M. M. W. B. (2014). Reflections on univariate and multivariate analysis of metabolomics data. *Metabolomics*, *10*(3), 361–374. https://doi.org/10.1007/s11306-013-0598-6

Sayo, H., & Hosokawa, M. (1980). Spin immunoassay of urinary testosterone (author's transl). *Yakugaku Zasshi*, *100*(1), 56–60.

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., Greiner, R., Sykes, B. D., & Marrie, T. J. (2007). Investigations of the effects of gender, diurnal variation, and age in human urinary metabolomic profiles. *Analytical Chemistry*, *79*(18), 6995–7004. https://doi.org/10.1021/ac0708588

Sugimoto, M., Saruta, J., Matsuki, C., To, M., Onuma, H., Kaneko, M., Soga, T., Tomita, M., & Tsukinoki, K. (2013). Physiological and environmental parameters associated with mass spectrometry-based salivary metabolomic profiles. *Metabolomics*, *9*(2), 454–463. https://doi.org/10.1007/s11306-012-0464-y

Suhre, K., Schwartz, J. E., Sharma, V. K., Chen, Q., Lee, J. R., Muthukumar, T., Dadhania, D. M., Ding, R., Ikle, D. N., Bridges, N. D., Williams, N. M., Kastenmüller, G., Karoly, E. D., Mohney, R. P., Abecassis, M., Friedewald, J., Knechtle, S. J., Becker, Y. T., Samstein, B., … Suthanthiran, M. (2016). Urine metabolite profiles predictive of human kidney allograft status. *Journal of the American Society of Nephrology*, *27*(2), 626–636. https://doi.org/10.1681/ASN.2015010107

Tarnopolsky, M. A., & Ruby, B. C. (2001). Sex differences in carbohydrate metabolism. *Current Opinion in Clinical Nutrition and Metabolic Care*, *4*(6), 521–526.

Thevenot, E. A., Roux, A., Xu, Y., Ezan, E., & Junot, C. (2015). Analysis of the human adult urinary metabolome variations with age, body mass index, and gender by implementing a comprehensive workflow for univariate and OPLS statistical analyses. *Journal of*

Proteome Research, *14*(8), 3322–3335. https://doi.org/10.1021/acs.jproteome.5b00354

Tippmann, S. (2015). Programming tools: Adventures with R. *Nature*, *517*(7532), 109–110. https://doi.org/10.1038/517109a

Torres-Benítez, A., Rivera-Montalvo, M., Sepúlveda, B., Castro, O., Nagles, E., Simirgiotis, M., García-Beltrán, O., & Areche, C. (2017). Metabolomic analysis of two *Parmotrema Lichens*: *P. robustum* (Degel.) Hale and *P. andinum* (Mull. Arg.) hale using UHPLC-ESI-OT-MS-MS. *Molecules*, *22*(11), 1861. https://doi.org/10.3390/molecules22111861

Vecht, C. J. (1993). Effect of age on treatment decisions in low-grade glioma. *Journal of Neurology, Neurosurgery and Psychiatry*, *56*(12), 1259–1264.

Wang, T. J., Larson, M. G., Vasan, R. S., Cheng, S., Rhee, E. P., McCabe, E., Lewis, G. D., Fox, C. S., Jacques, P. F., Fernandez, C., O'Donnell, C. J., Carr, S. A., Mootha, V. K., Florez, J. C., Souza, A., Melander, O., Clish, C. B., & Gerszten, R. E. (2011). Metabolite profiles and the risk of developing diabetes. *Nature Medicine*, *17*(4), 448–453. https://doi.org/10.1038/nm.2307

Weiss, R. H., & Kim, K. (2011). Metabolomics in the study of kidney diseases. *Nature Reviews Nephrology*, *8*(1), 22–33. https://doi.org/10.1038/nrneph.2011.152

Wild, C. P., Scalbert, A., & Herceg, Z. (2013). Measuring the exposome: A powerful basis for evaluating environmental exposures and cancer risk. *Environmental and Molecular Mutagenesis*, *54*(7), 480–499. https://doi.org/10.1002/em.21777

Wishart, D. S. (2016). Emerging applications of metabolomics in drug discovery and precision medicine. *Nature Reviews Drug Discovery*, *15*(7), 473–484. https://doi.org/10.1038/nrd.2016.32

Wishart, D. S., Feunang, Y. D., Marcu, A., Guo, A. C., Liang, K., Vázquez-Fresno, R., Sajed, T., Johnson, D., Li, C., Karu, N., Sayeeda, Z., Lo, E., Assempour, N., Berjanskii, M., Singhal, S., Arndt, D., Liang, Y., Badran, H., Grant, J., … Scalbert, A. (2018). HMDB 4.0: The human metabolome database for 2018. *Nucleic Acids Research*, *46*(D1), D608–D617. https://doi.org/10.1093/nar/gkx1089

Wismann, J., & Willoughby, D. (2006). Gender differences in carbohydrate metabolism and carbohydrate loading. *Journal of the International Society of Sports Nutrition*, *3*, 28–34. https://doi.org/10.1186/1550-2783-3-1-28

Wu, J., & Gao, Y. (2015). Physiological conditions can be reflected in human urine proteome and metabolome. *Expert Review of Proteomics*, *12*(6), 623–636. https://doi.org/10.1586/14789450.2015.1094380

Xia, J., Sinelnikov, I. V., Han, B., & Wishart, D. S. (2015). MetaboAnalyst 3.0–making metabolomics more meaningful. *Nucleic Acids Research*, *43*(W1), W251–W257. https://doi.org/10.1093/nar/gkv380

Zhang, A., Sun, H., Wu, X., & Wang, X. (2012). Urine metabolomics. *Clinica Chimica Acta*, *414*, 65–69. https://doi.org/10.1016/j.cca.2012.08.016

Zhang, A., Sun, H., Yan, G., Wang, P., & Wang, X. (2015). Metabolomics for biomarker discovery: Moving to the clinic. *BioMed Research International*, *2015*, 354671. https://doi.org/10.1155/2015/354671

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the Supporting Information section.

How to cite this article: Wang, T., Tang, L., Lin, R., He, D., Wu, Y., Zhang, Y., Yang, P., & He, J. (2021). Individual variability in human urinary metabolites identifies age-related, body mass index-related, and sex-related biomarkers. *Molecular Genetics & Genomic Medicine*, 9, e1738. https://doi.org/10.1002/mgg3.1738