Variability of Two Metabolomic Platforms in CKD

Eugene P. Rhee,1 Sushrut S. Waikar,2 Casey M. Rehholz,3,4 Zihe Zheng,5 Regis Perichon,6 Clary B. Clish,7 Anne M. Evans,6 Julian Avila,7 Michelle R. Denburg,8 Amanda Hyre Anderson,9 Ramachandran S. Vasan,9 Harold I. Feldman,5,10 Paul L. Kimmel,11,12 and Josef Coresh,1,4 on behalf of the CKD Biomarkers Consortium

Abstract

Background and objectives Nontargeted metabolomics can measure thousands of low-molecular-weight biochemicals, but important gaps limit its utility for biomarker discovery in CKD. These include the need to characterize technical and intraperson analyte variation, to pool data across platforms, and to outline analyte relationships with eGFR.

Design, setting, participants, & measurements Plasma samples from 49 individuals with CKD (eGFR < 60 ml/min per 1.73 m² and/or ≥1 g proteinuria) were examined from two study visits; 20 samples were repeated as blind replicates. To enable comparison across two nontargeted platforms, samples were profiled at Metabolon and the Broad Institute.

Results The Metabolon platform reported 837 known metabolites and 483 unnamed compounds (selected from 44,953 unknown ion features). The Broad Institute platform reported 594 known metabolites and 26,106 unknown ion features. A total of 381 known metabolites were shared across platforms (median correlation 0.89). Median CVs for day-to-day variability were 29.0% (Metabolon) and 24.9% (Broad Institute) for known metabolites, and 18.9% for (Metabolon) unnamed compounds and 24.5% for (Broad Institute) unknown ion features. Median coefficients of variation (CVs) across blind replicates were 14.6% (Metabolon) and 6.3% (Broad Institute) for known metabolites, and 18.9% for (Metabolon) unnamed compounds and 24.5% for (Broad Institute) unknown ion features.

Conclusions Nontargeted metabolomics quantifies >1000 analytes with low technical CVs, and agreement for overlapping metabolites across two leading platforms is excellent. Many metabolites demonstrate substantial intraperson variation and correlation with eGFR.


Introduction

Metabolomics refers to the systematic analysis of low-molecular-weight biochemicals in a biologic specimen (1). It is a promising tool for biomarker research in nephrology because kidney disease is associated with various metabolic disorders, because of the broad impact kidney function has on circulating metabolites, and because circulating metabolites may participate in disease pathogenesis (2,3). Studies to date have spanned individuals with normal kidney function to severe kidney failure, both acute and chronic (4). Several promising metabolites associated with onset of CKD or its progression have been highlighted, but most studies have been restricted to targeted methods that measure dozens to a few hundred metabolites (5–8).

Technologic improvements have led to the development of nontargeted methods that can measure several hundred to >1000 analytes in a biologic specimen, expanding the scope for discovery (9). Some studies have examined reproducibility of metabolomics platforms, both targeted and nontargeted (10–12), but a detailed analysis in the context of CKD is lacking (13). First, detailed knowledge is required on the technical and intraperson variation for all measurements, improving focus on higher-quality data. Second, approaches to comparing and pooling data across different platforms need to be developed, as their coverage of the metabolome can be non-overlapping. Finally, an assessment of how all metabolites correlate with GFR is essential for any application of these methods to nephrology research.

Here, we utilize two leading metabolomics platforms operated by Metabolon (M) and the Broad Institute (B), both with established track records of investigation in human biomarker research (14–19). We hypothesized that analyses of a common set of samples obtained from patients with CKD on both nontargeted platforms would permit characterization of intraassay, day-to-day, and interassay variation as well as characterize the relationship between GFR and the metabolome. Together, these studies provide a springboard for nontargeted metabolomics biomarker research and meta-analysis in CKD and beyond.
Materials and Methods

Study Design

We collected plasma samples from individuals with CKD attending a nephrology subspecialty practice at Brigham and Women’s Hospital (Boston, MA). Plasma samples from 49 individuals were examined from two study visits, with the evaluation of 20 samples repeated as blind replicates. All plasma samples, including blind replicates, were shipped to M and B in deidentified tubes. Metabolomics data generated by each platform were sent to Johns Hopkins University, where statistical analyses were performed.

Patients provided written informed consent to participate in this study, which was approved by the local institutional review board. Inclusion criteria included a diagnosis of CKD under the care of a nephrologist, defined as eGFR<60 ml/min per 1.73 m² with any degree of proteinuria or eGFR≥60 ml/min per 1.73 m² with ≥1 g proteinuria or albuminuria. Patients were excluded if they had a recent hospitalization or episode of AKI (>50% rise in serum creatinine over a 1 week period) within 3 months, active GN, reported or suspected urinary tract infection (defined as leukocytosis or bacteriuria with a rise in serum creatinine over a 1 week period) within 3 months, or a planned change by the attending nephrologist of the dose of a diuretic and/or antihypertensive medication during the study period.

Nontargeted Metabolomics

A detailed description of M and B nontargeted metabolomics methods are in the Supplemental Material. For both platforms, each experimental sample was analyzed using four distinct liquid chromatography-mass spectrometry (LC-MS) methods designed to cover a range of analyte polarity in both the positive and negative ion modes. Both platforms detected thousands of unknown ion features, from which they reported compounds of known identity (known metabolites). Among the unknown ion features, M reported a subset of these that are recurrent and reproducible within their data (unnamed compounds), whereas B reported all known ion features and unnamed compounds (B).

Other Measurements

Plasma and urine creatinine were measured using the Roche enzymatic method on a Roche COBAS 6000 chemistry analyzer. Plasma creatinin C and urinary albumin were measured on the Roche chemistry analyzer using a turbidimetric assay.

Statistical Analyses

Descriptive statistics were used to report the sociodemographic factors, clinical characteristics, and measures of kidney function for the study population. Because of the skewness of the metabolomics data, we used the coefficient of variation (CV)-ANOVA model recommended by Røraas et al. (20) to calculate CVs for blind duplicate pairs (n=20) and day-to-day variation pairs (n=49). For both analyses, we reported descriptive statistics for the Pearson correlation coefficients including mean, median, 5th percentile, 25th percentile, 75th percentile, 95th percentile, and proportion of analytes with a correlation coefficient >0.8. We also reported the median CV and the proportion of analytes with a CV <20% and <35% because the focus of discovery efforts will be the reliably measured metabolites. The assessment of technical and day-to-day variability was repeated among the subset of overlapping known metabolites detected by both platforms. Finally, we calculated the Pearson correlation between all analytes, after log transformation, and several estimates of kidney function: eGFR on the basis of serum creatinine (eGFRcr), serum cystatin C, and both serum creatinine and cystatin C. The distribution of correlation coefficients was presented in a kernel density plot for known metabolites and unnamed compounds (M) or unknown ion features (B).

Results

Study Sample and Data Overview

Characteristics of the 49 study participants are shown in Table 1, with mean age 58 years, diabetes prevalence 35%, and median eGFR 40.0 ml/min per 1.73 m² (interquartile range 25.0–52.0 ml/min per 1.73 m²). As shown in Supplemental Table 1, M reported 896 metabolites with known identity (known metabolites) and 488 compounds of unknown identity (unnamed compounds, selected from 44,953 detected unknown ion features). B reported 681 known metabolites and 28,630 deisotoped unknown ion features. After data cleaning, 837 (M) and 594 (B) known metabolites and 483 (M) unnamed compounds and 26,180 (B) unknown ion features were examined in subsequent analyses.

Technical Variation and Day-to-Day Variation for Known Metabolites

Technical variation was assessed using analysis of blind duplicates (n=20 pairs) (Figure 1A, Table 2). For known metabolites, correlations across blind duplicates was excellent for both platforms, with mean correlations of 0.83 (M) and 0.90 (B), and 634 (75.8%) (M) and 527 (88.7%) (B) metabolites with a correlation >0.8. Median CV was 14.6% (M) and 6.3% (B), and 557 (66.6%) (M) and 523 (88.1%) (B) metabolites had a CV<20%. Day-to-day variation was assessed using analysis of samples obtained from participants at different study visits (n=49 pairs) (Figure 1A, Table 2). For known metabolites, mean correlation across day-to-day variation pairs was 0.65 (M) and 0.66 (B), respectively, and 252 (30.1%) (M) and 166 (28.0%) (B) metabolites had a correlation >0.8 across visits. Median CV was 29.0% (M) and 24.9% (B), and 275 (32.9%) (M) and 241 (40.6%) (B) metabolites had a CV<20%. Data on technical and day-to-day variation for all known metabolites is shown in Supplemental Tables 2 and 3.

Technical Variation and Day-to-Day Variation for Unknown Compounds (M) and Unknown Ion Features (B)

Because both platforms are nontargeted, they produce data on thousands of ion features of unknown identity. As noted above, M reports on a subset of unnamed
Table 1. Clinical data and summary of filtration markers

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>All Subjects (n=49)</th>
<th>Subset for Blind Replicates (n=20)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Visit 1</td>
<td>Visit 2</td>
</tr>
<tr>
<td>Age, mean (SD)</td>
<td>58 (16)</td>
<td>63 (11)</td>
</tr>
<tr>
<td>Men, n %</td>
<td>27 (55)</td>
<td>12 (67)</td>
</tr>
<tr>
<td>Race, n %</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>30 (61)</td>
<td>12 (67)</td>
</tr>
<tr>
<td>Black</td>
<td>15 (31)</td>
<td>6 (33)</td>
</tr>
<tr>
<td>Hispanic</td>
<td>3 (6)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Asian</td>
<td>3 (2)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>eGFR, median (IQR)</td>
<td>40 (25–52)</td>
<td>40 (34–57)</td>
</tr>
<tr>
<td>eGFR&gt;60 ml/min per 1.73 m², n %</td>
<td>4 (8)</td>
<td>2 (11)</td>
</tr>
<tr>
<td>eGFR&lt;20 ml/min per 1.73 m², n %</td>
<td>11 (22)</td>
<td>3 (17)</td>
</tr>
<tr>
<td>Systolic BP, mm Hg, mean (SD)</td>
<td>133 (19)</td>
<td>130 (15)</td>
</tr>
<tr>
<td>Diastolic BP, mm Hg, mean (SD)</td>
<td>77 (14)</td>
<td>74 (13)</td>
</tr>
<tr>
<td>ACE or ARB, n %</td>
<td>35 (71)</td>
<td>12 (67)</td>
</tr>
<tr>
<td>Dihydrate, n %</td>
<td>20 (41)</td>
<td>8 (44)</td>
</tr>
<tr>
<td>ESKD, n %</td>
<td>4 (9)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Diabetes mellitus, n %</td>
<td>17 (35)</td>
<td>8 (44)</td>
</tr>
<tr>
<td>Diabetes mellitus and hypertension, n %</td>
<td>9 (53)</td>
<td>1 (5)</td>
</tr>
<tr>
<td>Plasma creatinine, mg/dl, median (IQR)</td>
<td>1.61 (1.30–2.23)</td>
<td>1.72 (1.34–2.32)</td>
</tr>
<tr>
<td>Plasma cystatin C, mg/L, median (IQR)</td>
<td>1.76 (1.42–2.53)</td>
<td>1.83 (1.42–2.45)</td>
</tr>
<tr>
<td>UACR, median (IQR)</td>
<td>0.14 (0.02–0.66)</td>
<td>0.13 (0.02–0.57)</td>
</tr>
</tbody>
</table>

IQR, interquartile range; ACE, angiotensin-converting enzyme inhibitor; ARB, angiotensin II receptor blocker; UACR, urine albumin-to-creatinine ratio.

*Multiply by 88.42 to convert to SI units (μmol/L).

compounds detected in plasma by the lab, rather than reporting all detected unknown features. Technical and day-to-day variation on reported unknowns (unnamed compounds for M, unknown ion features for B) are shown in Figure 1B and Table 3. Correlations across blind duplicates were excellent for M, with mean correlation 0.85, and modest for B, with mean correlation 0.53, yielding 256 unnamed compounds (M) and 11,344 unknown ion features (B). For overlapping, known metabolites, the median correlation of 0.89. The pathways represented by this overlap are shown in Supplemental Table 4, with lipid and amino acid metabolism most highly represented, followed by xenobiotics, nucleotides, vitamins and cofactors, and carbohydrates. Table 4 shows technical and day-to-day variation for the overlapping 381 known metabolites, demonstrating similar results across platforms. Technical performance for these overlapping metabolites was generally better than for other, nonoverlapping known metabolites.

Platform Overlap

Of the 837 M known metabolites and 593 B known metabolites, 381 were overlapping, and 113 of the overlapping known metabolites were identified by comparing B unknown ion features with M known metabolites, i.e., on the basis of mass alignment and correlations, followed by analysis of commercial standards. Figure 2 shows the distribution of correlation for these 381 known metabolites, with a median correlation of 0.89. The pathways represented by this overlap are shown in Supplemental Table 4, with lipid and amino acid metabolism most highly represented, followed by xenobiotics, nucleotides, vitamins and cofactors, and carbohydrates. Table 4 shows technical and day-to-day variation for the overlapping 381 known metabolites, demonstrating similar results across platforms. Technical performance for these overlapping metabolites was generally better than for other, nonoverlapping known metabolites.

Correlation with eGFR

Figure 3 depicts the correlation between metabolites levels and eGFRcr, including both known metabolites and unnamed compounds (M) or unknown ion features (B). For both M and B, many metabolites were correlated with eGFRcr at P<0.05, including 35.7% (M) and 18.9% (B) of known metabolites, and 23.1% of unnamed compounds (M) and 28.5% of unknown ion features (B). Results are similar for comparison with eGFR on the basis of serum cystatin C or serum creatinine and cystatin C (not shown). For overlapping, known metabolites, the median correlation with eGFRcr was −0.17 for both M and B, with interquartile range −0.43 to 0.07 for M and −0.45 to 0.06 for B. Correlations with eGFR for all metabolites are shown in Supplemental Tables 2 and 3, and the overlapping metabolites with a correlation <−0.8 with eGFR are shown in Supplemental Table 5.

Discussion

The measurement of select circulating metabolites, such as creatinine, urea, uric acid, lactic acid, glucose, triglycerides, and cholesterol, are integral to clinical medicine, with implications for diagnosis, prognosis, and treatment. Metabolomics studies seek to build on this paradigm, with the goal to identify novel markers and even functional participants in disease, including CKD and its complications. Nontargeted methods have expanded the potential scope for discovery, but have not been rigorously scrutinized as a tool for CKD biomarker discovery.

Unlike nucleic acids, which are limited to a handful of chemical motifs of similar abundance, metabolites span a variety of chemical classes with large differences in size and polarity, across a wide range of abundance. Thus, there is heterogeneity in how many and how well individual metabolites are measured. In our study, the M platform reported more known compounds, whereas the B platform
reported more unknown analytes. For known metabolites, the platforms demonstrated excellent repeatability, with both reporting >500 known metabolites with CV<20% across blind replicates. Importantly, the identity of known metabolites for both platforms has been confirmed with the use of reference standards, rather than relying on predictions on the basis of mass accuracy and comparison with databases (21).

Median repeatability was similarly excellent for M unnamed compounds, whereas it was modest for B unknown ion features. This relates in part to the fact that some of the >26,000 unknown ion features reported by the B platform are not unique metabolites, but rather ions derived from alternate ion products, organic solvents or tubing, or background noise. Recently, Mahieu and Patti (22) carefully examined nontargeted metabolomics data derived from Escherichia coli in culture. After accounting for alternate ion products and artifacts, followed by an analysis of 13C-fed bacterial cells (to confirm organic origin), these authors reduced 25,000 ion features in their data to <1000 unique analytes. The authors emphasize that the actual size of the metabolome in E. coli is larger and that their results simply outline what may be measured in a representative nontargeted analysis. The human metabolome, which incorporates endogenous metabolism, diet, and medications, and inputs from countless gut bacterial species is clearly larger. Further, these analyses in E. coli were performed on data from a single LC-MS method, in contrast to the four LC-MS runs utilized by the M and B platforms for this study.

Data reporting of unknown ion features weighs the risk of false discovery against maximizing the scope of data. In this study, the M data set focuses on the subset of ion features deemed to represent bona fide compounds (unnamed compounds), whereas the B data set reports a wider array of unknown ion features, acknowledging that unambiguous identifications requires subsequent collaboration and may be time-consuming or even unsuccessful. In either case, consideration of unnamed compounds or unknown ion features imposes an additional statistical penalty for multiple hypothesis testing. One approach for future studies would be to restrict analyses to the subset of unknown features with relatively better technical CVs, which would still yield ample opportunity for discovery (23). However, this approach would still require careful analysis of significant findings, given the possibility of redundant features such as adducts, multimers, and in-source fragments (24).

As with technical variability, heterogeneity in day-to-day metabolite variability is critical to understand for biomarker applications. Because circulating metabolites reflect both endogenous metabolism as well as exogenous inputs, some components of the metabolome change depending on fasting or feeding (25), as well as circadian oscillation (26). Substantial biologic variability does not necessarily rule a metabolite out as a potential biomarker or disease mediator. Indeed, blood glucose levels can change minute to minute, but fasting levels are still very useful clinically. For potential CKD biomarkers, however, low biologic...
variability is likely to be an asset (as with creatinine and cystatin C) (27). In one study of diabetic nephropathy and progression to ESKD, the authors only considered metabolites that had a correlation coefficient $>0.4$ on paired, fasting samples drawn 1–3 years apart, narrowing the number of metabolites considered from 262 to 119 (28). Here, we examined correlations across nonfasting samples to account for variable prandial status, as is frequently encountered in the clinic. In addition, we examined shorter term biologic variation, sampling individuals at visits approximately 2 weeks apart. Our results demonstrate substantial day-to-day variation in the plasma metabolome, with mean correlation across samples of approximately 0.6. Thus, mandating high correlation across paired measurements reduces the number of potential metabolite biomarkers.

In genomics, the ability to pool results across studies has been critical for replication and meta-analysis (29). Unlike sequencing data that is amenable to digital coding and meta-analysis, heterogeneity in how many and how well metabolites are measured across different platforms has made similar efforts in metabolomics uncommon. In one exception, a genetic association study of plasma metabolites, discovery was performed in one cohort profiled at B and validation was performed in another cohort profiled at M, yielding a total sample number of $>$3500 individuals (30). Thus, this study performed independent replication both in terms of study cohorts and analytical methods, enhancing confidence in the results. These analyses, however, were primarily anchored around abundant metabolites (e.g., histidine, phenylalanine, and xanthosine), and the data were generated using earlier, targeted iterations of the B and M methods.

Our examination of nontargeted metabolomics data from a common set of samples yields several observations. First, the overlapping known metabolites represent a subset of particularly well measured molecules, with generally better CVs across technical replicates than the other known metabolites. This group includes many metabolites abundant in blood, such as amino acids and lipids. Second, the number of overlapping known metabolites will continue to increase, building on the $>$100 added herein. As noted, the M platform also generates data across thousands of ion features. Although not reported, these can be extracted and analyzed for correlation in our data to generate additional metabolite identities. Use of additional LC-MS methods on both platforms (e.g., a dedicated complex lipids method at M not used in this study) would further enhance overlap. Third, agreement for the majority of overlapping measurements across platforms is strong, setting the stage for meta-analysis in future studies. Finally, there is also some nonoverlap in metabolome coverage across platforms.

### Table 2. Technical and day-to-day variation for known metabolites

<table>
<thead>
<tr>
<th>Correlations</th>
<th>CVs</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>Mean</td>
</tr>
<tr>
<td>Technical replicates (n=20)</td>
<td>Metabolon</td>
</tr>
<tr>
<td></td>
<td>Broad Institute</td>
</tr>
<tr>
<td>Day-to-day variation pairs (n=49)</td>
<td>Metabolon</td>
</tr>
<tr>
<td></td>
<td>Broad Institute</td>
</tr>
</tbody>
</table>

CV, coefficient of variation; $r$, Pearson correlation coefficient.

### Table 3. Technical and day-to-day variation for unnamed compounds (Metabolon) and unknown ion features (Broad Institute)

<table>
<thead>
<tr>
<th>Correlations</th>
<th>CVs</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>Mean</td>
</tr>
<tr>
<td>Technical replicates (n=20)</td>
<td>Metabolon</td>
</tr>
<tr>
<td></td>
<td>Broad Institute</td>
</tr>
<tr>
<td>Day-to-day variation pairs (n=49)</td>
<td>Metabolon</td>
</tr>
<tr>
<td></td>
<td>Broad Institute</td>
</tr>
</tbody>
</table>

CV, coefficient of variation; $r$, Pearson correlation coefficient.

*Selected from 44,953 detected unknown ion features.*
This reinforces an existing general challenge in metabolomics that no single analytic approach is truly comprehensive and that careful understanding of the methods utilized is crucial. For CKD biomarker research, the simultaneous deployment of different methodologies across different cohorts is prudent, to cast the widest net for discovery, and to minimize the risk inherent in selecting one analytical approach.

Kidney function affects the level of many circulating metabolites, an effect corroborated by our data as well as prior CKD metabolomics studies (31,32). Notably, inverse correlation with GFR does not necessarily indicate that a metabolite undergoes glomerular filtration, as the kidneys can also contribute to metabolite clearance through reabsorption, secretion, and metabolism. In addition, loss of kidney function can lead to alterations in diet, gut microbiota, insulin resistance, and other metabolic parameters that may influence the metabolome (33). Finally, a metabolite in theory could be inversely correlated with GFR because it causes kidney damage. Thus, statistical approaches may vary depending on the questions of interest. For biomarker studies, rigorous adjustment is required to identify specific signals with predictive power beyond GFR. By contrast, studies that seek to identify novel causal mediators may choose to examine unadjusted associations, followed by pathway analyses and mechanistic studies.

Strengths of this study include simultaneous assessment of several axes of metabolite variation, blinded metabolomics measurements, and independent statistical analysis. Several limitations also warrant mention. We restricted our study to two metabolomics platforms with established track records in human biomarker investigation. Other outstanding platforms are available, and as we have already noted, heterogeneity across methods means that no single approach is optimal. For practical purposes, we limited our analysis of day-to-day variation to two measurements over several weeks. A more in-depth study of repeated sampling would be required to disentangle the effects of diet, circadian oscillation, and other factors on biologic variation of the metabolome. Finally, although we focused on blood, a similar approach to urine is warranted (34,35).

Our results provide a framework for nontargeted metabolomics in CKD biomarker research. They exhibit the

<table>
<thead>
<tr>
<th>Table 4. Technical and day-to-day variation for overlapping known metabolites (n=381 metabolites)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Correlations</strong></td>
</tr>
<tr>
<td>Mean</td>
</tr>
<tr>
<td>Technical pairs (n=20)</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Day-to-day pairs (n=49)</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

CV, coefficient of variation; r, Pearson correlation coefficient.
ability of current methods to quantitate hundreds of known metabolites, and even more unknown ion features. They also outline the challenges imposed by technical and biologic variation, as well as correlation with GFR, that will need to be overcome if markers are to provide diagnostic or prognostic insights into kidney disease. Finally, as methods continue to improve, areas of overlap across platforms will no doubt expand, enhancing power to confirm the strongest biomarker associations.

Acknowledgments
This work is on behalf of the CKD Biomarkers Consortium, funded by the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) (grant U01 DK106981, Principal Investigator E.P.R.; grant U01 DK085689, Principal Investigator J.C.; grant U01 DK085660, Coprincipal Investigator S.S.W.). C.M.R. was supported by a mentored research scientist development award from the NIDDK (K01 DK107782).

Disclosures
R.P. and A.M.E. are employees of Metabolon, Inc. J.A. and C.B.C. are employees of the Broad Institute. E.P.R. is an Associate Member of the Broad Institute.

This article contains the following supplemental material online at http://cjasn.asnjournals.org/lookup/suppl/doi:10.2215/CJN.07070618/-/DCSupplemental.

Supplemental Material
Supplemental Table 1. Metabolite counts.
Supplemental Table 2. Technical variation, day-to-day variation, and correlation with eGFR for all Metabolon known metabolites.

Supplemental Table 3. Technical variation, day-to-day variation, and correlation with eGFR for all Broad Institute known metabolites.

Supplemental Table 4. Interplatform matched metabolite correlation, stratified by superpathways.

Supplemental Table 5. Shared metabolites with strong correlation with eGFR.

References


Received: June 13, 2018 Accepted: October 15, 2018

Published online ahead of print. Publication date available at www.cjasn.org.

**AFFILIATIONS**

1Nephrology Division and Endocrine Unit, Massachusetts General Hospital, Boston, Massachusetts; 2Renal Division, Brigham and Women’s Hospital, Boston, Massachusetts; 3Welch Center for Prevention, Epidemiology, and Clinical Research, Johns Hopkins University, Baltimore, Maryland; 4Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland; 5Department of Biostatistics, Epidemiology, and Informatics, 6Division of Nephrology, The Children’s Hospital of Philadelphia, and 7Department of Medicine, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania; 8Metabolon, Inc., Durham, North Carolina; 9Metabolite Profiling, Broad Institute of MIT and Harvard, Cambridge, Massachusetts; 10Sections of Preventive Medicine and Epidemiology and Cardiology, Department of Medicine, Boston University School of Medicine, Boston, Massachusetts; and 11Division of Kidney Urologic and Hematologic Diseases, National Institutes of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland