

METABOLOMICS REVEALS FIVE ENDOGENOUS BIOMARKERS IN HUMAN URINE AND PLASMA TO PREDICT CYP2D6 ACTIVITY

Gaëlle Magliocco¹, Alain Matthey¹, Nasim Bararpour², Timothée Joye², Yvonne Gloor¹, Jules Desmeules¹, Aurélien Thomas², and Youssef Daali¹

¹Division of Clinical Pharmacology and Toxicology, Geneva University Hospitals

²Forensic Toxicology and Chemistry Unit, CURML, Lausanne University Hospital, Geneva University Hospitals

January 22, 2021

Abstract

Background and Purpose: Individualized assessment of the activity of cytochrome P450 2D6 (CYP2D6), a highly variable drug-metabolizing enzyme, is performed through phenotyping during which a probe drug is administered to measure the enzyme's activity. In order to avoid any iatrogenic harm (allergic drug reaction, dosing error) related to the probe drug, the development of non-invasive tools for real-time phenotyping of CYP2D6 could significantly contribute to the expansion of precision medicine in clinical practice. This study focuses on the identification of endogenous markers of the CYP2D6 enzyme in human biofluids using a liquid chromatography (LC)-high-resolution mass spectrometry (HRMS)-based metabolomics approach. **Experimental Approach:** Data from a control session were compared to data from an inhibition session. Before the latter, healthy volunteers (extensive and ultrarapid metabolizers) received a daily dose of paroxetine 20 mg over seven days. CYP2D6 genotyping and phenotyping, using single oral dose of dextromethorphan 5 mg, were also performed in all participants. **Key Results:** In CYP2D6 extensive and ultrarapid metabolizers (n = 37), mean relative intensities of five features were significantly reduced during the inhibition session compared to the control session (fold changes [?] 0.67, FDR-adjusted P < 0.0001). Furthermore, mean relative intensities of these candidates were significantly higher in the CYP2D6 extensive-ultrarapid metabolizer group (n = 37) compared to the poor metabolizer group (n = 6) (fold changes [?] 0.67, P < 0.0001). **Conclusion and Implications:** The applied untargeted metabolomics strategy was able to identify five CYP2D6 endogenous metabolites, a promising discovery for non-invasive phenotyping and personalised medicine.

Gaëlle Magliocco (Orcid ID: 0000-0003-4352-5440)

RESEARCH ARTICLE

METABOLOMICS REVEALS FIVE ENDOGENOUS BIOMARKERS IN HUMAN URINE AND PLASMA TO PREDICT CYP2D6 ACTIVITY

Running title : CYP2D6 biomarkers discovery using untargeted metabolomics

Gaëlle Magliocco^{1,2}, Alain Matthey^{1,3}, Nasim Bararpour^{4,5}, Timothée Joye^{4,5}, Yvonne Gloor¹, Jules Desmeules^{1,2,3,6}, Aurélien Thomas^{4,5}*, Youssef Daali^{1,2,6}*

¹ Division of Clinical Pharmacology and Toxicology, Geneva University Hospitals, Geneva, Switzerland

² Institute of Pharmaceutical Sciences of Western Switzerland, University of Geneva, Geneva, Switzerland

³ Clinical Research Center, Geneva University Hospitals, Geneva, Switzerland

⁴ Forensic Toxicology and Chemistry Unit, CURML, Lausanne University Hospital, Geneva University Hospitals, Lausanne, Geneva, Switzerland

⁵ Faculty Unit of Toxicology, CURML, Faculty of Biology and Medicine, University of Lausanne, Lausanne Switzerland

⁶ Faculty of Medicine, University of Geneva, Geneva, Switzerland

* Co-Last authors

Corresponding author: Prof Youssef Daali, Division of Clinical Pharmacology and Toxicology, Geneva University Hospitals. Email: youssef.daali@hcuge.ch

Number of words (excluding Abstract, Methods, Reference and Figure Legends): 2878

Number of tables and figures: 10

Number of references: 49

Acknowledgments

The authors would like to thank the Clinical Research Center of Geneva University Hospital and the Genomics Platform of iGE3 for helpful contribution. This work was supported by the Development and Research Project of the Medical Direction of the Geneva University Hospitals.

Author contributions

G.M. wrote the manuscript, revised by all the authors. G.M, A.M, J.D., A.T. and Y.D. conceived and designed the research. G.M, A.M., N.B., T.J., Y.G. and J.D. performed the research. G.M., N.B., T.J., Y.G., A.T. and Y.D analyzed the data.

Conflict of interest/DISCLOSURE

The authors declare no conflict of interest.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request. Some data may not be made available because of privacy or ethical restrictions.

BULLET POINT SUMMARY

What is already known ?

CYP2D6 shows significant interindividual variability in activity due to genetic polymorphism and environmental causes.

Current CYP2D6 phenotyping methods require xenobiotics administration with risk of adverse events and time constraint.

What this study adds

Five endogenous urinary and plasma compounds were characterized as promising CYP2D6 biomarkers.

Mean relative intensities decreased significantly during drug-induced inhibition and in poor versus extensive metabolizers.

Clinical significance

Replacing CYP2D6 probe drugs with endogenous biomarkers is a step forward towards personalized medicine.

This methodology eliminates any risks due to exogenous substances and facilitates phenotyping in vulnerable populations.

ABSTRACT

Background and Purpose : Individualized assessment of the activity of cytochrome P450 2D6 (CYP2D6), a highly variable drug-metabolizing enzyme, is performed through phenotyping during which a probe drug is administered to measure the enzyme's activity. In order to avoid any iatrogenic harm (allergic drug reaction, dosing error) related to the probe drug, the development of non-invasive tools for real-time phenotyping of CYP2D6 could significantly contribute to the expansion of precision medicine in clinical practice. This study focuses on the identification of endogenous markers of the CYP2D6 enzyme in human biofluids using a liquid chromatography (LC)-high-resolution mass spectrometry (HRMS)-based metabolomics approach.

Experimental Approach : Data from a control session were compared to data from an inhibition session. Before the latter, healthy volunteers (extensive and ultrarapid metabolizers) received a daily dose of paroxetine 20 mg over seven days. CYP2D6 genotyping and phenotyping, using single oral dose of dextromethorphan 5 mg, were also performed in all participants.

Key Results : In CYP2D6 extensive and ultrarapid metabolizers (n = 37), mean relative intensities of five features were significantly reduced during the inhibition session compared to the control session (fold changes [?] 0.67, FDR-adjusted $P < 0.0001$). Furthermore, mean relative intensities of these candidates were significantly higher in the CYP2D6 extensive-ultrarapid metabolizer group (n = 37) compared to the poor metabolizer group (n = 6) (fold changes [?] 0.67, $P < 0.0001$).

Conclusion and Implications : The applied untargeted metabolomics strategy was able to identify five CYP2D6 endogenous metabolites, a promising discovery for non-invasive phenotyping and personalised medicine.

KEY WORDS

CYP450, phenotyping, metabolomics, CYP2D6, endobiotics

ABBREVIATIONS

AS Activity score

CYP2D6 Cytochrome P450 2D6

DDIs Drug-drug interactions

DEM Dextromethorphan

DOR Dextrorphan

EM Extensive metabolizer

PM Poor metabolizer

UM Ultrarapid metabolizer

UMR Urinary metabolic ratios

INTRODUCTION

Interindividual and intraindividual variability in drug response can lead to insufficient therapeutic efficacy or life-threatening adverse events (Kaddurah-Daouk et al., 2014). In this context, precision medicine aims to improve therapeutic outcomes by integrating the entire genetic and phenotypic knowledge specifically

related to an individual. Pharmacogenomics and pharmacometabolomics are both major and complementary approaches to precision medicine (Beger et al., 2016).

Pharmacogenomics is the use of patient-specific information associated with the genome to study individual response to drugs, while pharmacometabolomics focuses on the metabolome (profile of low molecular weight molecules within a biological system) (Schrimpe-Rutledge et al., 2016; Pang et al., 2019; Wake et al., 2019). Metabolomics allows identification and understanding of pathways involved in drug-response variation (Kaddurah-Daouk et al., 2014). It is also an important tool in the discovery of biomarkers that can be applied to personalized medicine (Villasenor et al., 2014; Jensen et al., 2017; Ivanisevic and Thomas, 2018; Yeung, 2018). Biomarkers help monitor the evolution of a disease and the corresponding response to drugs, as well as better predict the clinical outcomes (Kohler et al., 2017). For instance, testosterone glucuronide, when normalized by androsterone glucuronide, can be used as a urinary biomarker of an androgen- and drug-metabolizing enzyme (i.e. UGT2B17), as recently shown through targeted metabolomics analysis (Zhang et al., 2020). Five ω - and (ω -1)-hydroxylated medium-chain acylcarnitines have also been identified as novel CYP3A biomarkers using an untargeted metabolomics approach (Kim et al., 2018).

The cytochrome P450 2D6 (CYP2D6) is responsible for the metabolism of around 25% of all drugs used in clinical practice including antidepressants, analgesics, β -blocking agents and antipsychotics (Gaedigk, 2013). Prescribing CYP2D6 drug substrates is often challenging for physicians because of the large variability in the activity of this enzyme. CYP2D6 is a highly polymorphic gene locus and genotyping assays can be used to predict enzyme activity (Nofziger et al., 2020). However, relying only on genotyping has several limitations. First, it does not take into account environmental factors such as concomitant medications, food intake and disease-related factors (Gaedigk et al., 2018). Second, depending on the technology and database used, some of the rare variants may not be screened or even identified, and an allele may be erroneously categorized as functional (Gaedigk et al., 2018). And third, when duplication or multiplication is detected, a majority of copy number tests do not distinguish which of the two allele has several copies (Langae et al., 2015; Shah et al., 2016). Therefore, in clinical practice, precision medicine must rely on both real-time phenotyping and genotyping in order to provide the best possible recommendations. Currently, CYP2D6 phenotyping requires the administration of an exogenous probe drug specifically metabolised by this isoenzyme (Samer et al., 2013; Magliocco et al., 2019). Microdosing of the probe drug and enhanced detection capacities of mass spectrometry have lowered the risk of probe-related side effects. However, potential iatrogenic harm (allergic reaction, dosing errors) would only be totally eliminated if endogenous probes were available (Magliocco et al., 2019; Magliocco and Daali, 2020). A recent review summarized human endogenous compounds that have been tagged as potential CYP2D6 markers (Magliocco et al., 2019). One of them stands out. It is a very promising urinary biomarker named M1 (m/z 444.3102). It was characterised, but not yet structurally identified in a non-targeted metabolomics study (Tay-Sontheimer et al., 2014). Some *in vitro* and animal studies have also demonstrated that CYP2D6 metabolizes the endocannabinoid anandamide (Snider et al., 2008).

Our main objective in this study was to explore the presence of CYP2D6 biomarkers in human urine and plasma, using an untargeted metabolomics approach. For this purpose, healthy volunteers were invited to two sessions (control vs inhibitory). Prior to the inhibitory session, volunteers received over 7 days, a daily dose (10 or 20 mg) of the strong CYP2D6 inhibitor paroxetine. The CYP2D6 genotype and phenotype were also integrated in the data analysis.

METHODS

Study design and population

This study protocol (NCT04188028) was approved by the Geneva Research Ethics Committee and the Swiss Agency for Therapeutic Products (Swissmedic). All participants provided written informed consent before inclusion. Protocol conception and trial conduct were performed in accordance with the Declaration

of Helsinki ethical principles and the Good Clinical Practice guidelines of the International Congress of Harmonization.

Inclusion criteria were the following: age between 18-65 years, body mass index between 18-27 kg m⁻², CYP2D6 genotype activity score (AS) = 0 or [?] 1 (Gaedigk et al., 2008), reliable contraception during the whole study, including a barrier method.

Exclusion criteria included: pregnancy/breastfeeding, any pathology, drug or food affecting CYP activity, tobacco consumption ([?] 10 cigarettes/day), alcohol intake 2 days prior to session 1 and during paroxetine intake, hepatic impairment, medical history of chronic alcoholism or abuse of psychoactive drugs, regular use of psychotropic substances, drug sensitivity, psychiatric disorders, and Beck Score [?] 10 (question related to suicide >0).

Forty-three healthy volunteers participated in this study, which was conducted in two sessions. Each session included the oral administration of 5 mg dextromethorphan (DEM) (BEXIN syrup, Spirig Healthcare, Egerkingen, Switzerland) to participants after an overnight fast and urine collection for 4 hours following the administration of DEM for CYP2D6 phenotyping. For metabolomic analyses, prior to DEM administration, urine samples were also collected over a full 24-hour period and venous blood samples were collected in tubes containing EDTA (BD Vacutainer, Plymouth, UK) immediately before DEM ingestion. Breakfast was served 1 hour after DEM intake. At session 2, the study course was similar but participants were asked to take 20 mg (10 mg for poor metabolizer (PM) subjects) of paroxetine (PAROXETIN-MEPHA, Basel, Switzerland), a time-dependant inhibitor, every morning for one week (7 doses in total) with the breakfast (Storelli et al., 2019). Participants were specifically asked about the time at which paroxetine tablets were taken and were asked to bring back empty blister packs to verify compliance. For women participating in the study, a pregnancy test was performed at inclusion and at each session prior to any medication administration. Plasma was obtained through centrifugation at 2,750g for 10 min. All blood and urine samples were stored at -80°C until analysis.

Quantification of dextromethorphan and dextrorphan

Subsequent to chemical hydrolysis and liquid-liquid extraction (Daali et al., 2008), DEM and dextrorphan (DOR) were quantified in urine by liquid chromatography-tandem mass spectrometry (Sciex, Darmstadt, Germany). CYP2D6 phenotype was determined based on the urinary metabolic ratio dextromethorphan to dextrorphan ($UMR_{DEM/DOR}$) as follows: PM phenotype ($UMR_{DEM/DOR}$ [?] 0.3), intermediate metabolizer (IM) phenotype ($UMR_{DEM/DOR}$ between 0.03-0.3), extensive metabolizer (EM) phenotype ($UMR_{DEM/DOR}$ between 0.003-0.3) and ultrarapid metabolizer (UM) phenotype ($UMR_{DEM/DOR}$ < 0.003) (Gaedigk et al., 2008).

CYP450 genotyping

Genomic DNA was extracted from whole blood (200 µl) using the QIAamp DNA Blood Mini Kit (Qia-gen, Hombrechtikon, Switzerland). Fourteen CYP2D6 allelic variants were screened using the TaqMan® OpenArray® PGx Panel (Thermo Fisher Scientific, Waltham, USA) performed on the QuantStudio 12K Flex real-time PCR system in compliance with the manufacturer's instructions. The following mutations were considered: 2850C>T, 4180G>C, 2549delA (*3), 100C>T (*4, *10), 1846G>A (*4A), 1707delT (*6), 2935A>C (*7), 1758G>T (*8), 2613_2615delAGA(*9), 124G>A (*12), 1758G>A (*14), 1023C>T (*17), 3183G>A (*29), 2988G>A (*41). Regarding CYP2C9 the following mutations were measured: 430C>T, 3608C>T (*2), 1075A>C, 42614A>C (*3). CYP2C19*2 (681G>A, 19154G>A), CYP2C19*3 (636G>A, 17948G>A) and CYP2C19*17 (806C>T) were also determined, as well as CYP3A4*22 (15389C>T) and CYP3A5*3 (6986A>G).

CYP2D6 Taqman(r) Copy Number Assay (assay ID: Hs00010001_cn targeting exon 9, Applied Biosystems, Foster City, USA) was performed on a 7900HT Fast Real-Time PCR System (Applied Biosystems, Thermo

Fisher Scientific, CA, USA) instrument for the detection of gene deletion (*5 allele) and duplication.

CYP2C9, CYP2C19, CYP2D6 and CYP3A activity scores were assigned using previously developed scoring system (Gaedigk et al., 2008; Elens et al., 2011; Karnes et al., 2020; Lima et al., 2020). Values of 0, 0.5, 1 and 1.5, were assigned to the non-functional, reduced function, fully functional and gain-of-function alleles, respectively (Tay-Sontheimer et al., 2014). In the case of CYP2D6, the values for alleles with two or more gene copies are multiplied by the number of gene (Gaedigk et al., 2018). Summing the values of the two alleles gives the AS of a genotype (Gaedigk et al., 2018).

Untargeted metabolomics analysis by LC-HRMS

300 μ L of methanol/ethanol (50:50) containing hydrocodone-D6 and phenobarbital-D5 100 ng/mL (internal standards for positive and negative modes, respectively) were added to 100 μ L of urine or plasma for protein precipitation. Samples were centrifuged for 20 min at 16,000g. The supernatant was then evaporated under a stream of nitrogen, reconstituted in 100 μ L of 10% methanol and 5 μ L were injected into the LC-MS system.

Non-targeted metabolomics analyses were carried out using a LC system Ultimate 3000 coupled to a Q Exactive Plus system (Thermo Scientific Fisher, Bremen, Germany). (Forchelet et al., 2018; Kowalczyk et al., 2018) Separation was performed with a Kinetex[®] C18 column (50 \times 2.1 mm, 2.6 μ m) from Phenomenex (Brebühler, Switzerland) with mobile phases consisting of water (A) and methanol (B) both containing 0.1% formic acid. The flow rate was fixed at 0.3 mL/min over 13 minutes. Gradient program was set as follows: 2% B (0-0.3 minutes), 2-98% B (0.3-6 minutes), 98-100% B (6-9 minutes), 100-2% B (9-9.1 minutes), and 2% B (9.1-13 minutes). Quality controls (i.e. pooled aliquots of all clinical study samples) were included in the analytical sequence at regular intervals. Data was acquired in a full scan mode in both positive and negative polarities. The parameters were set as follows: the capillary voltage at 3.2 kV and 2.5 kV in positive and negative mode, respectively, sheath and auxiliary gas flow rate at 40 and 10 respectively, capillary temperature at 320 degC and S-lens RF level at 50.

Untargeted metabolomics Data and Statistical Analysis

The raw UPLC-HRMS files were converted to .mzXML format using MSConvert (ProteoWizard 3.0, <http://proteowizard.sourceforge.net/>) and pre-processed using the XCMS Online platform for features detection, chromatogram alignment, isotope annotation and data visualization (<https://xcmsonline.scripps.edu>).

All data transformation and statistical analyses were performed using MetaboAnalyst (<https://www.metaboanalyst.ca/>). Data were sum-normalized, Pareto-scaled and log-transformed. Subsequently, features were filtered, and only those with a CV less than 20% in the QC samples were selected. Isotopes were filtered out and finally (Kim et al., 2018), ions of zero intensity in >20% of all participants in both sessions were excluded.

Zero values were replaced by the half of the minimum value found for the corresponding hit (Xia and Wishart, 2011). Principal Component Analysis (PCA) was performed using QC samples to assess performance and stability of the system. Data Volcano plots were generated in order to filter metabolites that displayed both significant fold changes ([?] 1.5 or [?] 0.67) and statistical significance (FDR adjusted P -value < 0.05) between the control and the inhibition session in non-PM subjects ($n = 37$). The significant features obtained were then filtered out according to genotype: fold changes of relative intensity in the CYP2D6 EM-UM group ($n = 37$) compared to the PM group ($n = 6$) [?] 0.67 or [?] 1.50 (P -value < 0.05). The data and statistical analysis comply with recommendations of the British Journal of Pharmacology on experimental design and analysis (Curtis et al., 2018).

CYP2D6 biomarkers identification

Metabolites molecular formulas were investigated further using MS fragmentation and isotope pattern analysis with SIRIUS 4.0.1 (Duhrop et al., 2019). The glucuronide metabolites were enzymatically deconjugated prior to MS fragmentation using the β -*Glucuronidase / Arylsulfatase* mixture from Roche Diagnostics (Mannheim, Germany) in sodium acetate buffer (1 M, pH 5.0) at 37°C (Schmidt et al., 2013). Main metabolomics databases: LIPID MAPS® (<https://www.lipidmaps.org/>), METLIN (<https://metlin.scripps.edu/>) and HMDB (<http://www.hmdb.ca/>) were then used to assist in the identification of molecular structures of significant features on the basis of the available experimental data (i.e. exact molecular weights, molecular formulas, fragmentation patterns).

Production Reaction Monitoring analysis

To improve sensitivity, validate and refine results, a semi-quantitative method using HRMS-based PRM was developed. The chromatographic separation was performed using a LC system Vanquish coupled to a Q Exactive Focus system (Thermo Scientific, Bremen, Germany). The preparation of urine and plasma samples as well as the chromatographic and mass spectrometry conditions were identical to those of the metabolomics analyses, except that the extracts were concentrated twice (reconstitution in 50 μ L of 10% methanol). Hydrocodone-d6 at 15 ng/mL was used as internal standard. The resolution was set at 17'500 for the fragmentation experiments with an AGC target of 5e4 and a maximum IT of 100 ms. The NCE values for each compound were set individually and urinary creatinine concentration was used for data normalization.

Production Reaction Monitoring Data and Statistical analysis

Comparisons of two dependent and independent groups were performed using paired and unpaired t test (two-tailed), respectively. Measures of associations were established using Spearman's rank correlation. The statistical analyses were performed using GraphPad Prism 8.0.1 software (San Diego, USA). A *P* value below 0.05 was considered statistically significant. The data and statistical analysis comply with recommendations of the British Journal of Pharmacology on experimental design and analysis (Curtis et al., 2018).

Nomenclature of Targets and Ligands

Key protein targets and ligands in this article are hyperlinked to corresponding entries in <http://www.guidetopharmacology.org>, the common portal for data from the IUPHAR/BPS Guide to PHARMACOLOGY (Harding et al., 2018), and are permanently archived in the Concise Guide to PHARMACOLOGY 2019/20 (Alexander et al., 2019).

RESULTS

Healthy subjects

A total of 43 healthy subjects were enrolled (**Table 1**). Forty-two completed the study while one among the PM subjects only attended the control session. The mean age was 24 (range 19-29) with a slightly higher proportion of females (55.8 %, n = 24). Among women participants, 45.8% used oral contraceptive pill (n = 11).

CYP2D6 genotype and phenotype

Based on genotyping analyses, 6 participants were classified as PM subjects (genetic-predicted activity score (gAS) = 0), 33 as EM subjects (1 [?] gAS [?] 2) and 4 as UM subjects (gAS > 2) (Gaedigk et al., 2008; Crews et al., 2014). UMR_{DEM/DOR} were measured to establish CYP2D6 phenotype. As illustrated in **Figure 1a**, mean CYP2D6 activity was significantly reduced after paroxetine intake compared to control session (*P* <

0.0001), demonstrating an effective inhibition of CYP2D6 activity by paroxetine. In addition, a significant Spearman’s rank correlation coefficient of $r_s = 0.791$ ($P < 0.0001$) was found between the logarithm of $UMR_{DEM/DOR}$ and AS groups (**Figure 1b**).

Untargeted metabolomics analysis

Using untargeted metabolomics assays, the goal of this project was to identify biomarkers of CYP2D6 in urine and plasma reflecting the activity of the enzyme. **Figure 2** shows the flowchart of the data analysis, from data extraction to statistical analysis. After the filtering steps, 8926 and 5997 ions in plasma and urine, respectively, were processed for statistical analysis, including sum-normalisation, log-transformation and Pareto scaling. PCA scores plot revealed a tight clustering of QC samples (**Supplementary Figure S1**) indicating high experimental quality for both urine and plasma samples.

As seen in **Table 2**, five endogenous metabolites were significantly decreased in urine and/or plasma during the CYP2D6 inhibition phase relative to baseline with the following m/z in positive mode: 220.1545, 416.3159, 432.3108, 444.3108 and 597.3382 (595.3236 in negative mode). The largest reduction was observed for m/z 597.3382 (0.13-fold). In parallel, their intensity was significantly lower in PM subjects than in EM-UM volunteers (fold changes [?] 0.67). These observations strongly imply that these five features are metabolites produced via the CYP2D6 enzyme.

Structure identification

All MS/MS fragmentation patterns are shown in **Figure 3**. Due to the presence of a glucuronide moiety, the feature m/z 597.3382 was also enzymatically hydrolysed prior to MS/MS fragmentation, resulting in m/z 421.3061 (-176.0321 Da).

MS/MS fragmentation of m/z 220.1543 was difficult to obtain because several compounds with close mass (i.e. m/z 220.0966, 220.1329) co-elute. At 10 eV it is however possible to observe four losses of water (-18.0109 Da): 202.1436, 184.1331, 166.1226 and 148.1120.

The compounds m/z 416.3159, 432.3108 and 444.3108 have identical fragmentation patterns, with a major fragment at 98.0967. The skeletal structure appears therefore similar for these three compounds, which differ by one or two atoms: carbon loss between 444.3108 and 432.3108 (-12.0000 Da), oxygen loss between 432.3108 and 416.3159 (-15.9949 Da), and carbon-oxygen loss between 444.3108 and 416.3159 (-27.9949 Da). Regarding the feature m/z 444.3108, we observed the same MS/MS fingerprint described by Tay-Sontheimer et al. (fragment ions at m/z : 98.0967, 370.2733, 206.1900, 56.0501, 55.0549, 150.1275 and 81.0703) (Tay-Sontheimer et al., 2014).

As described above, one of the major fragments of m/z 597.3382 is 421.3061, which corresponds to neutral loss of a glucuronide moiety. Chromatograms before and after hydrolysis shown in **Supplementary Figure S2** confirm the presence of a glucuronide since the conjugated peak (m/z 597.3382) decreases and the deconjugated peak (m/z 421.3061) increases after hydrolysis.

Molecular formula for all compounds was obtained through fragmentation and isotope pattern analysis using SIRIUS 4.0.1. Results are described in **Table 3**. Fragmentation trees are presented in **Supplementary Figure S3-S8**. Interestingly, all the features contain one or two nitrogen atoms. Using exact mass, mass spectral databases were employed for potential features annotation and identification as shown in **Table 3**. Results were, however, inconclusive in HMDB (no results for the query masses). In LIPID MAPS, only m/z 416.3159 showed a potential match: N-linoleyl dopamine. Nonetheless, this lipid normally has a characteristic and predominant fragment at 137, which was not observed, making this finding unlikely (Thomas et al., 2009). In METLIN, the feature m/z 444.3108 matches a prostaglandin derivative: 17-phenyl trinor PGF2 α diethyl amide. However, MS/MS fragmentation patterns are not concordant.

Based on these findings and the physicochemical properties of these compounds, in particular their high retention times and molecular weight, it is likely that they belong to lipid species, including a lipid glucuronide.

Relative quantification using a PRM method

MS parameters were first optimized to enhance data quality using parallel reaction monitoring (PRM) mode. In order to obtain the optimal collision energy (CE) for each metabolite, the NCE were considered individually using a step of 10. The results are presented in **Table 4**. All analytes were detectable in the urine samples of EM and UM subjects, with the exception of PM volunteers regarding m/z 432.3108 and 444.3108. Metabolites corresponding to m/z 220.1543 and 432.3108 were not measurable in any of the plasma samples. Also, metabolites with m/z 416.3159 and 444.3108 were not detectable in plasma samples of PM participants.

The results confirmed the significant down-regulation observed after paroxetine intake compared to the control session for all five hits in both urine and plasma samples (**Figure 4a** and **5a**). Likewise, the compounds were not detectable or significantly down-regulated in PM subjects compared to EM-UM participants (**Figure 4b** and **5b**). Going even further, significant correlations were observed between $\log(\text{area})$ and $\log(\text{UMR}_{\text{DEM/DOR}})$ (**Figure 4c** and **5c**). Significant correlations between $\log(\text{area})$ and CYP2D6 activity score were also described, with UM subjects ($\text{gAS} > 2$) having the highest mean intensities (**Figure 4d** and **5d**). All P were < 0.05 .

Mean relative intensity of m/z 220.1543, 416.3159 and 597.3382 was unchanged in PM subjects after paroxetine intake compared to baseline, indicating that changes are solely due to CYP2D6 inhibition (**Figure 6**). Indeed, PM individuals do not express the enzyme CYP2D6. Therefore, a down-regulation in PM subjects would indicate the presence of false positives rather than changes due to CYP2D6 inhibition.

No significant correlation was observed with any CYP450 activity score other than CYP2D6 (**Supplementary Table S1**).

DISCUSSION

This study analysed the metabolome of 43 healthy volunteers in order to identify novel endogenous biomarkers associated with CYP2D6 activity, using non-targeted metabolomics assays applied to urine and plasma samples. Drug-induced CYP2D6 inhibition and the search for CYP2D6 polymorphisms were taken into account for the identification of endogenous metabolites. Five endogenous metabolites could potentially act as probe for testing CYP2D6 activity as shown in **Table 2**. All of them were significantly reduced after seven days of paroxetine intake compared to baseline. These results were reinforced and validated thanks to the significant reduction observed for the five candidates in the PM group compared to the EM-UM group. One of them had a similar mass-to-charge ratio to the urinary biomarker M1 (m/z 444.3102) reported earlier (Tay-Sontheimer et al., 2014). In their study, M1 was below detection limit in PM paediatric subjects and its levels were reduced by drug-induced inhibition of CYP2D6 in adults. However, only urine samples were tested and we hypothesise that their methodology might not have been sensitive enough to detect the four other compounds highlighted in our study since they have lower signals than m/z 444.3108. Until now, they did not characterize the chemical structure of m/z 444.3108. The formal identification of metabolites is one of the main challenges of untargeted metabolomics (Alonso et al., 2015; Ivanisevic and Thomas, 2018). Multiple databases have been developed and are regularly updated to help scientists with this process. But a large proportion of the compounds, which have been revealed by untargeted metabolomic profiling, remains unidentified (Blaženović et al., 2018). For structure elucidation, fragmentation mass spectra is required to improve confidence in metabolite identification (e.g. matching of an experimental MS/MS spectrum with a reference fragmentation spectrum) (Alonso et al., 2015). Analysis of isotope patterns is another mean for the determination of the most likely elemental composition of metabolites. SIRIUS 4 is a software that combines both isotope pattern and MS/MS data through fragmentation tree (Dührkop et al., 2019). It was

successfully used within the framework of this project, allowing the elucidation of molecular formulas. The physicochemical properties of the five hits show that they are most likely lipids, one of them being a lipid glucuronide. More precisely, since they all contain a nitrogen element, they could potentially belong to fatty amides class. Interestingly, anandamide, a fatty acid amide that belongs to the class of endocannabinoids, is known to be metabolized by the CYP2D6 enzyme into 20-hydroxyeicosatetraenoic acid ethanolamide as well as 5,6-, 8,9-, 11,12- and 14,15- epoxyeicosatrienoic acid ethanolamides as demonstrated through in vitro experiments using recombinant CYP2D6 (Farrell and Merkler, 2008; Snider et al., 2008). Furthermore, it is generally known that a majority of CYP2D6 substrates are lipophilic bases with a protonable nitrogen atom (Ingelman-Sundberg, 2005). In this work, metabolomics databases (i.e. METLIN and LIPID MAPS) revealed the following matches: N-linoleoyl dopamine for m/z 416.3159 and 17-phenyl trinor PGF 2α diethyl amide for m/z 444.3108, two fatty acid amides compounds. However, MS/MS fragmentation patterns were not concordant. More in-depth analyses are required to achieve a detailed structure elucidation since the monoisotopic masses and MS/MS fragmentation patterns of the five hits appear to be unknown from the metabolomics databases used. Complete structure elucidation will require complementary analytical methods such as nuclear magnetic resonance (NMR) after extraction and purification of samples through preparative HPLC (Dias et al., 2016). Conveniently, all biomarker candidates are present in the urine. Large volumes of urine are relatively easy to obtain, which makes it an optimal starting material for purification and concentration for NMR structure identification (Whiley et al., 2019).

The targeted assay in PRM mode confirmed the results obtained from the untargeted LC-MS based metabolomics approach. It showed a significant decrease in the intensity of the features intensity after paroxetine intake compared to baseline. It is worth noting that signals from UM volunteers are often increased rather than decreased during the inhibition phase compared to baseline especially for m/z 220.1543, 416.3159, 432.3108 and 444.3108 as seen in **Figure 4a and 5a**. In addition of being an inhibitor of CYP2D6, paroxetine is also a substrate for this enzyme. As demonstrated in numerous studies, UM subjects have lower concentrations of paroxetine than EM patients (Hicks et al., 2015). CYP2D6 inhibition may therefore be less strong in such participants and may account for inconsistent results. However, the $UMR_{DEM/DOR}$ does not display such pattern. It is then likely that the identification of the structure of endogenous substrates would allow refining these results by measuring the ratios of endogenous substrate/metabolite.

Even when using a more sensitive MS-based PRM experiment, we were still unable to detect signals in PM subjects for the features m/z 416.3159 in plasma, 432.3108 in urine and 444.3108 in both. The same observation was reported for m/z 444.3108 (Tay-Sontheimer et al., 2014). Therefore, if certain metabolites are absent among this population group, it may explain differences in behaviour (e.g. impulsivity, anxiety) and disease susceptibility observed between PM and other individuals (Peñas-Lledó and LLerena, 2014). Additionally, $\log(\text{area}/\text{creatinine})$ in urine and $\log(\text{area})$ in plasma of the 5 endogenous compounds were both significantly correlated with CYP2D6 AS and $\log(UMR_{DEM/DOR})$. This confirms the potential ability of these compounds to accurately predict CYP2D6 activity. The best correlation with CYP2D6 AS was observed for m/z 597.3382 ($r_s = 0.710$) in plasma (**Figure 5d**). It is worth noting that the Spearman's rank correlation coefficients measured in this study for all five hits were always lower than that observed for $UMR_{DEM/DOR}$ ($r_s = 0.791$) (**Figure 1**). As previously explained, more significant correlations are expected once metabolites and substrates are fully identified, since measurement of the metabolic ratio between a specific substrate and a metabolite could be achieved similarly to $UMR_{DEM/DOR}$. Additionally, the use of a normalized ratio (substrate/metabolite), could correct for potentially highly variable metabolite concentrations due to the influence of different factors of variability, such as circadian rhythm, diet, physiological or pathophysiological conditions. As an example, urinary excretion of 6β -hydroxycortisol, a CYP3A metabolite, varies considerably throughout the day, showing a strong diurnal rhythm. However, when normalized to cortisol, the variation is no longer significant within the day (Lee, 1995). Similarly, plasma 4β -hydroxycholesterol, another CYP3A metabolite, is ideally normalized to cholesterol in order to provide more robust data (Mao et al., 2017; Aklillu et al., 2020).

Once the couple metabolite/substrate is fully identified and quantifiable, specificity for CYP2D6 enzyme should be assessed using, for example, recombinant enzyme assays in order to exclude the influence of other

metabolic pathways. In this study, no correlation with CYP2C9, CYP2C19 and CYP3A genotypes were observed with the identified biomarkers. In addition, no impact of oral contraceptives (moderate CYP1A2 and CYP3A inhibitor) (Samer et al., 2013) was observed on mean relative intensities of the five endogenous CYP2D6 metabolites (data not shown). But these results provide only some preliminary insights into the CYP2D6 specificity and further *in vitro* studies should be performed to fully validate specificity of the metabolic pathway.

In conclusion, non-targeted metabolomics enabled the identification of five potential endogenous CYP2D6 metabolites presumably related structurally to the class of fatty amides, including a glucuronide compound. Each of these candidate biomarkers could map the functionality of this enzyme. Further studies will focus on complete structure elucidation using complementary analytical method such as NMR. Once identified and validated, noninvasive prediction of CYP2D6 activity based on these candidates could greatly improve current phenotyping strategies by being readily available at any time and completely bypassing the need of administering exogenous components and thus the risk of adverse events.

REFERENCES

- Aklillu, E., Zumla, A., Habtewold, A., Amogne, W., Makonnen, E., Yimer, G., et al. (2020). Early or deferred initiation of efavirenz during rifampicin-based TB therapy has no significant effect on CYP3A induction in TB-HIV infected patients. *Br. J. Pharmacol.* *n/a* :
- Alexander, S.P.H., Christopoulos, A., Davenport, A.P., Kelly, E., Mathie, A., Peters, J.A., et al. (2019). THE CONCISE GUIDE TO PHARMACOLOGY 2019/20: G protein-coupled receptors. *Br. J. Pharmacol.* *176* : S21–S141.
- Alonso, A., Marsal, S., and Julià, A. (2015). Analytical Methods in Untargeted Metabolomics: State of the Art in 2015. *Front. Bioeng. Biotechnol.* *3* :
- Beger, R.D., Dunn, W., Schmidt, M.A., Gross, S.S., Kirwan, J.A., Cascante, M., et al. (2016). Metabolomics enables precision medicine: ‘A White Paper, Community Perspective’. *Metabolomics Off. J. Metabolomic Soc.* *12* : 149.
- Blaženović, I., Kind, T., Ji, J., and Fiehn, O. (2018). Software Tools and Approaches for Compound Identification of LC-MS/MS Data in Metabolomics. *Metabolites* *8* : 31.
- Crews, K.R., Gaedigk, A., Dunnenberger, H.M., Leeder, J.S., Klein, T.E., Caudle, K.E., et al. (2014). Clinical Pharmacogenetics Implementation Consortium guidelines for cytochrome P450 2D6 genotype and codeine therapy: 2014 update. *Clin. Pharmacol. Ther.* *95* : 376–382.
- Curtis, M.J., Alexander, S., Cirino, G., Docherty, J.R., George, C.H., Gienbycz, M.A., et al. (2018). Experimental design and analysis and their reporting II: updated and simplified guidance for authors and peer reviewers. *Br. J. Pharmacol.* *175* : 987–993.
- Daali, Y., Cherkaoui, S., Doffey-Lazeyras, F., Dayer, P., and Desmeules, J.A. (2008). Development and validation of a chemical hydrolysis method for dextromethorphan and dextrophan determination in urine samples: Application to the assessment of CYP2D6 activity in fibromyalgia patients. *J. Chromatogr. B* *861* : 56–63.
- Dias, D.A., Jones, O.A.H., Beale, D.J., Boughton, B.A., Benheim, D., Kouremenos, K.A., et al. (2016). Current and Future Perspectives on the Structural Identification of Small Molecules in Biological Systems. *Metabolites* *6* : 46.
- Dührkop, K., Fleischauer, M., Ludwig, M., Aksenov, A.A., Melnik, A.V., Meusel, M., et al. (2019). SIRIUS 4: a rapid tool for turning tandem mass spectra into metabolite structure information. *Nat. Methods* *16* : 299–302.
- Elens, L., Bouamar, R., Hesselink, D.A., Haufroid, V., Heiden, I.P. van der, Gelder, T. van, et al. (2011). A new functional CYP3A4 intron 6 polymorphism significantly affects tacrolimus pharmacokinetics in kidney

transplant recipients. *Clin. Chem.* *57* : 1574–1583.

Farrell, E.K., and Merkler, D.J. (2008). Biosynthesis, degradation, and pharmacological importance of the fatty acid amides. *Drug Discov. Today* *13* : 558–568.

Forchelet, D., Béguin, S., Sajic, T., Bararpour, N., Pataky, Z., Frias, M., et al. (2018). Separation of blood microsamples by exploiting sedimentation at the microscale. *Sci. Rep.* *8* : 14101.

Gaedigk, A. (2013). Complexities of CYP2D6 gene analysis and interpretation. *Int. Rev. Psychiatry Abingdon Engl.* *25* : 534–553.

Gaedigk, A., Dinh, J.C., Jeong, H., Prasad, B., and Leeder, J.S. (2018). Ten Years' Experience with the CYP2D6 Activity Score: A Perspective on Future Investigations to Improve Clinical Predictions for Precision Therapeutics. *J. Pers. Med.* *8* : 15.

Gaedigk, A., Simon, S.D., Pearce, R.E., Bradford, L.D., Kennedy, M.J., and Leeder, J.S. (2008). The CYP2D6 activity score: translating genotype information into a qualitative measure of phenotype. *Clin. Pharmacol. Ther.* *83* : 234–242.

Harding, S.D., Sharman, J.L., Faccenda, E., Southan, C., Pawson, A.J., Ireland, S., et al. (2018). The IUPHAR/BPS Guide to PHARMACOLOGY in 2018: updates and expansion to encompass the new guide to IMMUNOPHARMACOLOGY. *Nucleic Acids Res.* *46* : D1091–D1106.

Hicks, J., Bishop, J., Sangkuhl, K., Müller, D., Ji, Y., Leckband, S., et al. (2015). Clinical Pharmacogenetics Implementation Consortium (CPIC) Guideline for CYP2D6 and CYP2C19 Genotypes and Dosing of Selective Serotonin Reuptake Inhibitors. *Clin. Pharmacol. Ther.* *98* : 127–134.

Ingelman-Sundberg, M. (2005). Genetic polymorphisms of cytochrome P450 2D6 (CYP2D6): clinical consequences, evolutionary aspects and functional diversity. *Pharmacogenomics J.* *5* : 6–13.

Ivanisevic, J., and Thomas, A. (2018). Metabolomics as a Tool to Understand Pathophysiological Processes. *Methods Mol. Biol. Clifton NJ* *1730* : 3–28.

Jensen, B.C., Parry, T.L., Huang, W., Beak, J.Y., Ilaiwy, A., Bain, J.R., et al. (2017). Effects of the kinase inhibitor sorafenib on heart, muscle, liver and plasma metabolism in vivo using non-targeted metabolomics analysis. *Br. J. Pharmacol.* *174* : 4797–4811.

Kaddurah-Daouk, R., Weinshilboum, R.M., and Pharmacometabolomics Research Network (2014). Pharmacometabolomics: implications for clinical pharmacology and systems pharmacology. *Clin. Pharmacol. Ther.* *95* : 154–167.

Karnes, J.H., Rettie, A.E., Somogyi, A.A., Huddart, R., Fohner, A.E., Formea, C.M., et al. (2020). Clinical Pharmacogenetics Implementation Consortium (CPIC) Guideline for CYP2C9 and HLA-B Genotypes and Phenytoin Dosing: 2020 Update. *Clin. Pharmacol. Ther.* *n/a* :

Kim, B., Lee, J., Shin, K.-H., Lee, S., Yu, K.-S., Jang, I.-J., et al. (2018). Identification of ω - or (ω -1)-Hydroxylated Medium-Chain Acylcarnitines as Novel Urinary Biomarkers for CYP3A Activity. *Clin. Pharmacol. Ther.* *103* : 879–887.

Kohler, I., Hankemeier, T., Graaf, P.H. van der, Knibbe, C.A.J., and Hasselt, J.G.C. van (2017). Integrating clinical metabolomics-based biomarker discovery and clinical pharmacology to enable precision medicine. *Eur. J. Pharm. Sci. Off. J. Eur. Fed. Pharm. Sci.* *109S* : S15–S21.

Kowalczyk, L., Matet, A., Dor, M., Bararpour, N., Daruich, A., Dirani, A., et al. (2018). Proteome and Metabolome of Subretinal Fluid in Central Serous Chorioretinopathy and Rhegmatogenous Retinal Detachment: A Pilot Case Study. *Transl. Vis. Sci. Technol.* *7* : 3.

Langae, T., Hamadeh, I., Chapman, A.B., Gums, J.G., and Johnson, J.A. (2015). A Novel Simple Method for Determining CYP2D6 Gene Copy Number and Identifying Allele(s) with Duplication/Multiplication.

PLOS ONE 10 : e0113808.

Lee, C. (1995). Urinary 6 β P-hydroxycortisol in humans: Analysis, biological variations, and reference ranges. *Clin. Biochem.* 28 : 49–54.

Lima, J.J., Thomas, C.D., Barbarino, J., Desta, Z., Van Driest, S.L., El Rouby, N., et al. (2020). Clinical Pharmacogenetics Implementation Consortium (CPIC) Guideline for CYP2C19 and Proton Pump Inhibitor Dosing. *Clin. Pharmacol. Ther.*

Magliocco, G., and Daali, Y. (2020). Modern approaches for the phenotyping of cytochrome P450 enzymes in children. *Expert Rev. Clin. Pharmacol.* 0 : 1–4.

Magliocco, G., Thomas, A., Desmeules, J., and Daali, Y. (2019). Phenotyping of Human CYP450 Enzymes by Endobiotics: Current Knowledge and Methodological Approaches. *Clin. Pharmacokinet.* 58 : 1373–1391.

Mao, J., Martin, I., McLeod, J., Nolan, G., Horn, R. van, Vourvahis, M., et al. (2017). Perspective: 4 β -hydroxycholesterol as an emerging endogenous biomarker of hepatic CYP3A. *Drug Metab. Rev.* 49 : 18–34.

Nofziger, C., Turner, A.J., Sangkuhl, K., Whirl-Carrillo, M., Agúndez, J.A.G., Black, J.L., et al. (2020). PharmVar GeneFocus: CYP2D6. *Clin. Pharmacol. Ther.* 107 : 154–170.

Pang, H., Jia, W., and Hu, Z. (2019). Emerging Applications of Metabolomics in Clinical Pharmacology. *Clin. Pharmacol. Ther.* 106 : 544–556.

Peñas-Lledó, E.M., and LLerena, A. (2014). CYP2D6 variation, behaviour and psychopathology: implications for pharmacogenomics-guided clinical trials. *Br. J. Clin. Pharmacol.* 77 : 673–683.

Samer, C.F., Lorenzini, K.I., Rollason, V., Daali, Y., and Desmeules, J.A. (2013). Applications of CYP450 testing in the clinical setting. *Mol. Diagn. Ther.* 17 : 165–184.

Schmidt, L., Müller, J., and Göen, T. (2013). Simultaneous monitoring of seven phenolic metabolites of endocrine disrupting compounds (EDC) in human urine using gas chromatography with tandem mass spectrometry. *Anal. Bioanal. Chem.* 405 : 2019–2029.

Schrimpe-Rutledge, A.C., Codreanu, S.G., Sherrod, S.D., and McLean, J.A. (2016). Untargeted Metabolomics Strategies-Challenges and Emerging Directions. *J. Am. Soc. Mass Spectrom.* 27 : 1897–1905.

Shah, R.R., Gaedigk, A., LLerena, A., Eichelbaum, M., Stingl, J., and Smith, R.L. (2016). CYP450 genotype and pharmacogenetic association studies: a critical appraisal. *Pharmacogenomics* 17 : 259–275.

Snider, N.T., Sikora, M.J., Sridar, C., Feuerstein, T.J., Rae, J.M., and Hollenberg, P.F. (2008). The endocannabinoid anandamide is a substrate for the human polymorphic cytochrome P450 2D6. *J. Pharmacol. Exp. Ther.* 327 : 538–545.

Storelli, F., Desmeules, J., and Daali, Y. (2019). Genotype-sensitive reversible and time-dependent CYP2D6 inhibition in human liver microsomes. *Basic Clin. Pharmacol. Toxicol.* 124 : 170–180.

Tay-Sontheimer, J., Shireman, L.M., Beyer, R.P., Senn, T., Witten, D., Pearce, R.E., et al. (2014). Detection of an endogenous urinary biomarker associated with CYP2D6 activity using global metabolomics. *Pharmacogenomics* 15 : 1947–1962.

Thomas, A., Hopfgartner, G., Giroud, C., and Staub, C. (2009). Quantitative and qualitative profiling of endocannabinoids in human plasma using a triple quadrupole linear ion trap mass spectrometer with liquid chromatography. *Rapid Commun. Mass Spectrom. RCM* 23 : 629–638.

Villaseñor, A., Ramamoorthy, A., Silva dos Santos, M., Lorenzo, M.P., Laje, G., Zarate, C., et al. (2014). A pilot study of plasma metabolomic patterns from patients treated with ketamine for bipolar depression: evidence for a response-related difference in mitochondrial networks. *Br. J. Pharmacol.* 171 : 2230–2242.

Wake, D.T., Ilbawi, N., Dunnenberger, H.M., and Hulick, P.J. (2019). Pharmacogenomics: Prescribing Precisely. *Med. Clin. North Am.* *103* : 977–990.

Whiley, L., Chekmeneva, E., Berry, D.J., Jiménez, B., Yuen, A.H.Y., Salam, A., et al. (2019). Systematic Isolation and Structure Elucidation of Urinary Metabolites Optimized for the Analytical-Scale Molecular Profiling Laboratory. *Anal. Chem.* *91* : 8873–8882.

Xia, J., and Wishart, D.S. (2011). Web-based inference of biological patterns, functions and pathways from metabolomic data using MetaboAnalyst. *Nat. Protoc.* *6* : 743–760.

Yeung, P.K. (2018). Metabolomics and Biomarkers for Drug Discovery. *Metabolites* *8* : 11.

Zhang, H., Basit, A., Wolford, C., Chen, K.-F., Gaedigk, A., Lin, Y.S., et al. (2020). Normalized Testosterone Glucuronide as a Potential Urinary Biomarker for Highly Variable UGT2B17 in Children 7–18 Years. *Clin. Pharmacol. Ther.* *107* : 1149–1158.

TableS

Table 1. Demographic characteristics of participants

Parameters	All subjects (n = 43)
Age (years), mean (range)	24 (19-29)
BMI (kg m ⁻²), mean (range)	22 (19-27)
Female, n (%)	24 (56)
Oral contraceptive intake, n (%)	11 (26)
CYP2D6 allele frequencies, %	
1 (wild-type)	41
2	27
4	17
5	3
6	1
9	1
10	3
17	1
41	5
CYP2C9 allele frequencies, %	
1 (wild-type)	80
2	9
3	10
CYP2C19 allele frequencies, %	
1 (wild-type)	65
2	13
17	22
CYP3A4 allele frequencies, %	
1 (wild-type)	95
22	5
CYP3A5 allele frequencies, %	
1 (wild-type)	12
3	88

BMI, body mass index

Table 2. Significant hits in urine and plasma obtained from Volcano plots, including mean fold-changes of normalized intensity in the inhibition session versus control session, and mean fold-change of normalized

intensity in the PM subjects versus EM-UM subjects.

m/z	RT (min)	Ionization Mode	Fold-change	P -value (adjusted)
Plasma	Plasma	Plasma	Plasma	Plasma
416.3159	4.16	+	0.60	4.48×10^{-4}
595.3236	5.25	-	0.16	4.10×10^{-5}
597.3382	5.25	+	0.13	9.38×10^{-8}
Urine	Urine	Urine	Urine	Urine
220.1543	2.51	+	0.54	2.63×10^{-7}
416.3159	4.16	+	0.67	4.53×10^{-3}
432.3108	4.11	+	0.56	1.39×10^{-2}
444.3108	4.35	+	0.67	3.50×10^{-5}

Table 3. Exact mass, molecular formula and adducts obtained using SIRIUS 4.0.1, and results obtained using METLIN and LIPID MAPS databases

Monoisotopic mass	Molecular Formula	Adduct	Potential METLIN assignment	Potential LIPID MAPS assignment
219.1471	$C_{10}H_{21}NO_4$	[M+H] ⁺	NA	NA
415.3086	$C_{26}H_{41}NO_3$	[M+H] ⁺	NA	N-linoleoyl dopamine
431.3036	$C_{26}H_{41}NO_4$	[M+H] ⁺	NA	NA
443.3036	$C_{27}H_{41}NO_4$	[M+H] ⁺	17-phenyl trinor PGF2 α diethyl amide	NA
596.3309	$C_{30}H_{48}N_2O_{10}$	[M+H] ⁺	NA	NA
420.2988	$C_{24}H_{40}N_2O_4$	[M+H] ⁺	NA	NA

Table 4. MS parameters in PRM mode of the five target analytes and their detectability in urine and plasma samples

Analyte	Precursor Ion, m/z	Product Ion, m/z	NCE, eV	RT (min)	Detection urine	Detection plasma
Unknown	220.1543	202.1436	10	2.25	Yes	No
Unknown	416.3159	98.0967	65	3.90	Yes	Yes*
Unknown	432.3108	98.0967	65	3.84	Yes*	No
Unknown	444.3108	98.0967	65	4.09	Yes*	Yes*
Unknown	597.3382	84.0812	50	4.96	Yes	Yes

*Except in PM samples

Figure legends

Figure 1. (a) 4-hour dextromethorphan/dextrorphan individual urinary metabolic ratios ($UMR_{DEM/DOR}$) at baseline (control session) and after inhibition by paroxetine (inhibition session), following intake of dextromethorphan 5 mg per os ($P < 0.0001$). Ultrarapid metabolizers ($gAS > 2$, $n = 4$) are shown in red and extensive metabolizers ($1 \leq gAS \leq 2$, $n = 33$) in blue. Error bars represent the standard deviations of log-transformed $UMR_{DEM/DOR}$. (b) Correlation between 4-hour log-transformed dextromethorphan/dextrorphan urinary metabolic ratios ($\log(UMR_{DEM/DOR})$) and the CYP2D6 activity score ($n = 43$) with whiskers indicating the 10th and 90th percentiles ($P < 0.0001$)

Figure 2. Flowchart of the non-targeted metabolomics approach to identify biomarkers reflecting CYP2D6

activity. N represents the number of metabolic features after each step with plasma biomarkers in red and urinary biomarkers in yellow

Figure 3. Targeted Selected Ion Monitoring/data-dependent-MS² of all features of interest in ESI+ mode

Figure 4. Log(area/creatinine) in **urine** off all CY2D6 biomarkers identified in metabolomics analyses measured with product ion monitoring **(a)** before and after paroxetine intake, including means and standard deviations on each side. Ultrarapid metabolizers ($n = 4$) are shown in red and extensive metabolizers ($n = 33$) in blue. **(b)** EM-UM subjects ($n = 37$) versus PM subjects ($n = 6$) with whiskers indicating the 10th and 90th percentiles. **(c)** Correlation with log(UMR_{DEM/DOR}). Control session ($n = 43$) is represented by square and inhibition session ($n = 42$) by triangle. **(d)** Correlation with CYP2D6 activity score score ($n = 43$) with whiskers indicating the 10th and 90th percentiles. All $P < 0.0001$

Figure 5 . Log(area) in plasma off all CY2D6 biomarkers identified in metabolomics analyses measured with product ion monitoring **(a)** before and after paroxetine intake, including means and standard deviations on each side. Ultrarapid metabolizers ($n = 4$) are shown in red and extensive metabolizers ($n = 33$) in blue. **(b)** EM-UM subjects ($n = 37$) versus PM subjects ($n = 6$) with whiskers indicating the 10th and 90th percentiles. **(c)** Correlation with log(UMR_{DEM/DOR}). Control session ($n = 43$) is represented by square and inhibition session ($n = 42$) by triangle. **(d)** Correlation with CYP2D6 activity score ($n = 43$) with whiskers indicating the 10th and 90th percentiles. All $P < 0.0001$

Figure 6. Log(area/creatinine) or log(area) at baseline (control session) and after inhibition by paroxetine (inhibition session) ($n = 5$). On each side means values and standard deviations.

Hosted file

Table 1.pdf available at <https://authorea.com/users/391135/articles/505317-metabolomics-reveals-five-endogenous-biomarkers-in-human-urine-and-plasma-to-predict-cyp2d6-activity>

Hosted file

Table 2.pdf available at <https://authorea.com/users/391135/articles/505317-metabolomics-reveals-five-endogenous-biomarkers-in-human-urine-and-plasma-to-predict-cyp2d6-activity>

Hosted file

Table 3.pdf available at <https://authorea.com/users/391135/articles/505317-metabolomics-reveals-five-endogenous-biomarkers-in-human-urine-and-plasma-to-predict-cyp2d6-activity>

Hosted file

Table 4.pdf available at <https://authorea.com/users/391135/articles/505317-metabolomics-reveals-five-endogenous-biomarkers-in-human-urine-and-plasma-to-predict-cyp2d6-activity>







