Electronic Supplementary Information

^{1}H-NMR based metabolomics study for the detection of the human urine metabolic profile effects of *Origanum Dictamnus* tea ingestion.

Panteleimon G. Takis[*†‡], Mariam-Eleni Oraiopoulou[†+], Constantinos Konidaris[†] and Anastassios N. Troganis*[†]*

[†]Department of Biological Applications and Technology, University of Ioannina, GR-451 10 Ioannina, Greece; and [‡]Department of Chemistry, University of Ioannina, GR-451 10 Ioannina, Greece

[‡]Present address: Giotto Biotech S.r.l., Via Madonna del Piano 6, 50019 Sesto Fiorentino, Firenze, Italy

[+]Present address: Computational Bio-Medicine Laboratory (CBML), Institute of Computer Science (ICS), Foundation for Research & Technology – Hellas (ICS-FORTH), 100 N. Plastira St., GR-70013, Heraklion, Crete, Greece

*To whom correspondence should be addressed:

Dr. Panteleimon G. Takis; E-mail: ptakis@cc.uoi.gr

and

Assoc. Prof. Anastassios N. Troganis; E-mail: atrogani@uoi.gr

tel.: +30-26-5100-8851 and fax: +30-26-5100-8889

One-way ANOVA analysis*

One-way ANOVA is a univariate method, which calculates the variance of each variable without considering the dependency of the variables on each other (Brownlee, K. A. Statistical Theory and Methodology in Science and Engineering Wiley: New York, 1960.). The sum of squares of the distance between a value xi and its average describes the total variance (TSS):

$$TSS = \sum_i (x_i - \bar{x})^2 \qquad (1)$$

TSS is equal to the between-group ($SS_{between}$) and the within-group (SS_{within}) variations:

$$TSS = SS_{between} + SS_{within} \qquad (2)$$

The SS variations depend on the number of observations in each group, so, in order to evaluate more accurately the relative amplitudes of these variations, the mean squares are calculated by dividing each variation by its degree of freedom. Consequently, the mean between ($MS_{between}$) and within group (MS_{within}) variations and the *F*-statistics by Fisher-Snedector rule are calculated, where the latter is compared to a theoretical value from Fisher tables defined by the chosen *p*-value (here $p = 0.05$):

$$MS_{between} = \frac{SS_{between}}{k-1} \qquad (3)$$

$$MS_{within} = \frac{SS_{within}}{n-k} \qquad (4)$$

$$F = \frac{MS_{between}}{MS_{within}} \qquad (5)$$

where *k* and *n* are the number of groups (2 in our case: with and without Dictamnus tea ingestion) and the number of individuals (maximum 8 in our case), respectively.

For each *F* statistics, there is a corresponding *p*-value, which is the probability to observe an effect of the same amplitude under the null hypothesis. p-values are computed for each variable of the data matrix X (buckets) with respect to the descriptive variables contained in the Y matrix (with and without dictamnus ingestion phases). This kind of analysis is fully implemented in the statistical toolbox of Matlab platform.

*B. J. Blaise, L. Shintu, B. Elena, L. Emsley, M.-E. Dumas and P. Toulhoat, *Anal. Chem.*, 2009, **81**, 6242–6251.

^{1}H-NMR Urine profile

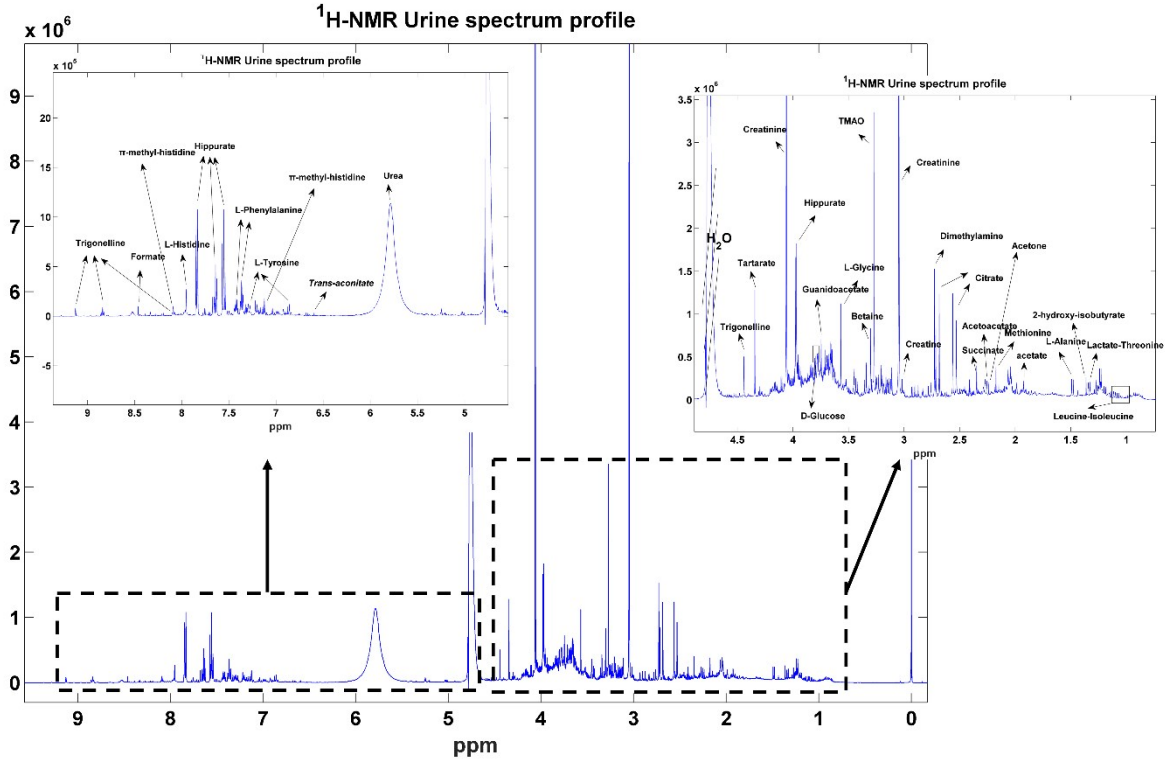


Figure S1. Several metabolites assignment in a ^{1}H NMR urine spectrum. Several metabolites (including those that are analyzed in our study) are annotated on the aromatic and the rest region of the urine ^{1}H NMR spectrum.

Guanidoacetate assignment via spiking experiments

(examples for the subjects which exhibited guanidoacetate variations)

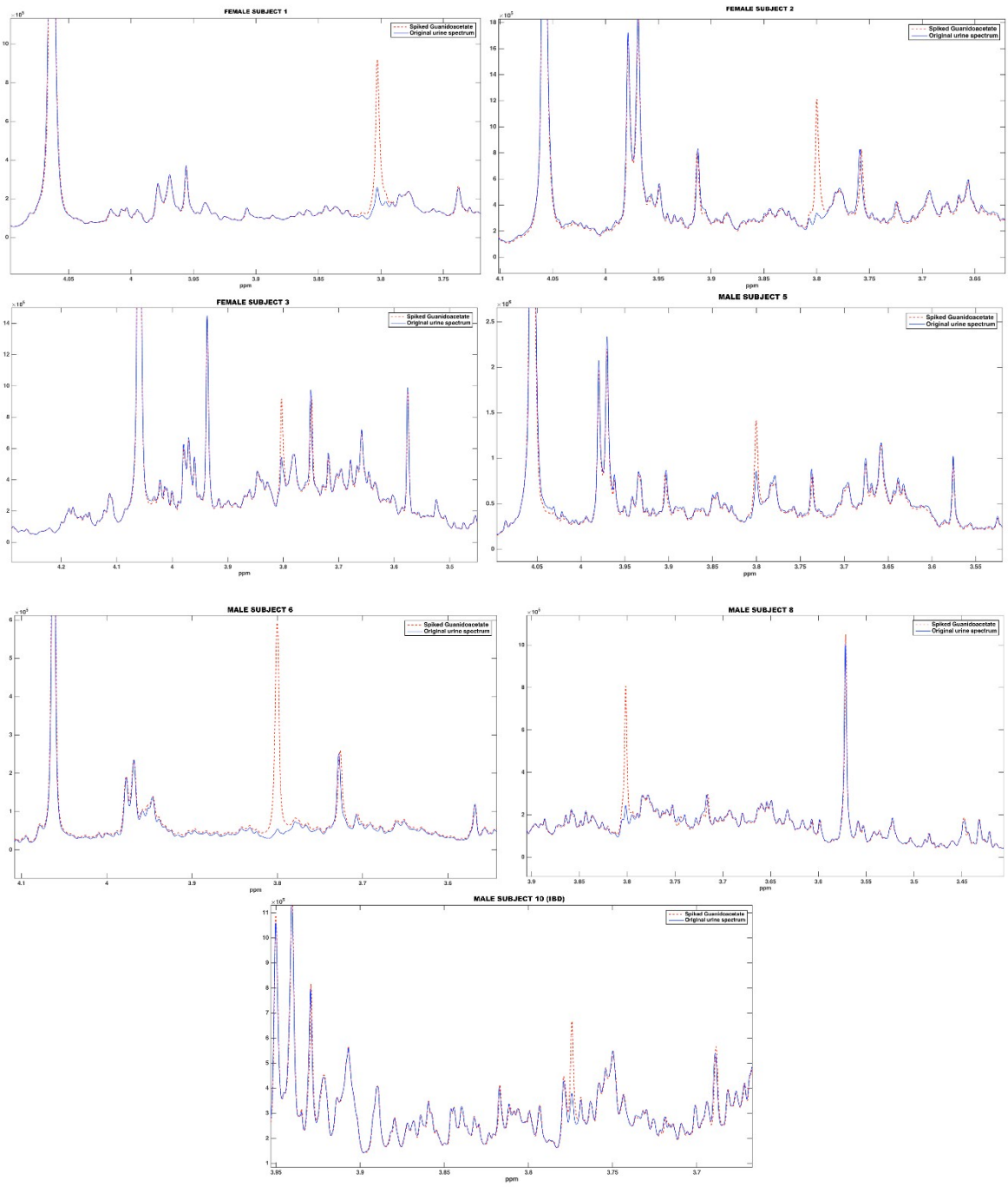
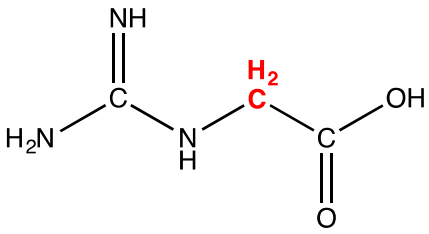


Figure S2. Spiking ^{1}H NMR experiments with guanidoacetate metabolite (red dashed elevated intensities of a singlet ^{1}H NMR peak). As depicted even without spiking the ^{1}H NMR peak of guanidoacetate (singlet, resonating within 3.76-3.80 ppm in our urine spectra) is quite distinctive for assignment. Here 7 examples of spiking are presented applied to the subjects' urine samples that guanidoacetate concentration levels were statistical significantly perturbed from dictamnus tea ingestion.

Table S1. Several 2D-NMR HSQC (^{13}C–^{1}H) spectra for guanidoacetate's assignment. In particular, the red font ^{13}C and its 2 attached protons NMR chemical shifts (δ) in ppm are reported.

Urine Samples HSQC (^{13}C–^{1}H) 2D-NMR spectra	**Guanidoacetate** 	
	δ ^{13}C (ppm)	**δ ^{1}H (ppm)**
1	47.53	3.8074
2	47.33	3.7863
3	47.50	3.8065
4	47.57	3.8101
5	47.45	3.7908
6	47.23	3.7698
7	47.50	3.8002

Probabilistic Quotient Normalization statistical analysis result

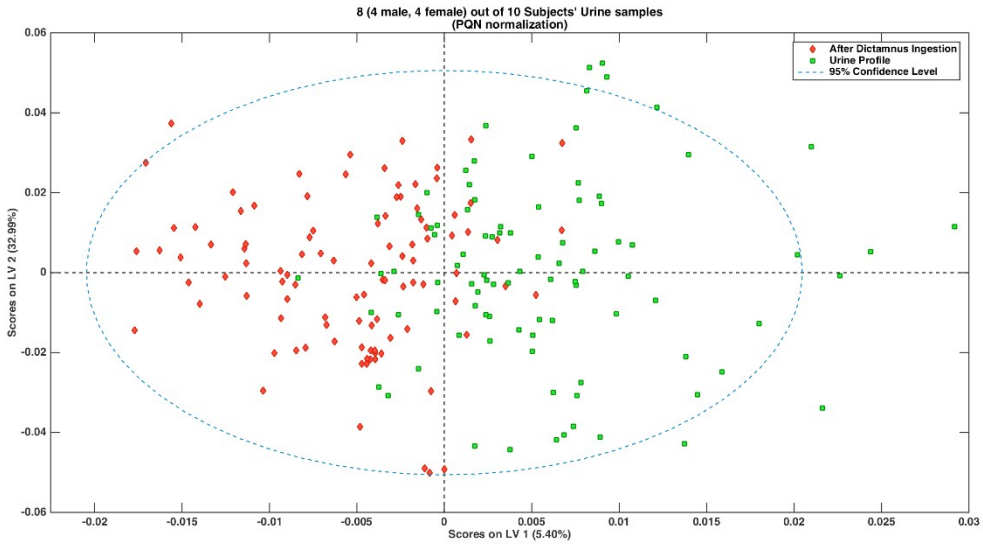


Figure S3. Score plot of the OPLS-DA classification of 200 urine samples ^{1}H-NMR spectra between 8 healthy subjects before (100 samples) and during dictamnus tea (100 samples) consumption after applying Probabilistic Quotient Normalization (PQN) of the urine NMR datasets.

Table S2. The cross-validated (CV) values of sensitivity, specificity and accuracy of the subjects' urine profile without and after dictamnus ingestion in the OPLS-DA derived model based upon their urine samples' ^{1}H-NMR spectra, after Probabilistic Quotient Normalization (PQN).

Urine samples ^{1}H-NMR 1D-noesy spectra	**After Dictamnus Ingestion**	**Urine Profile (without Dictamnus)**
Sensitivity (CV)	82.7 %	79.6 %
Specificity (CV)	79.6 %	82.7 %
Accuracy (CV)	82.1 %	81.2 %

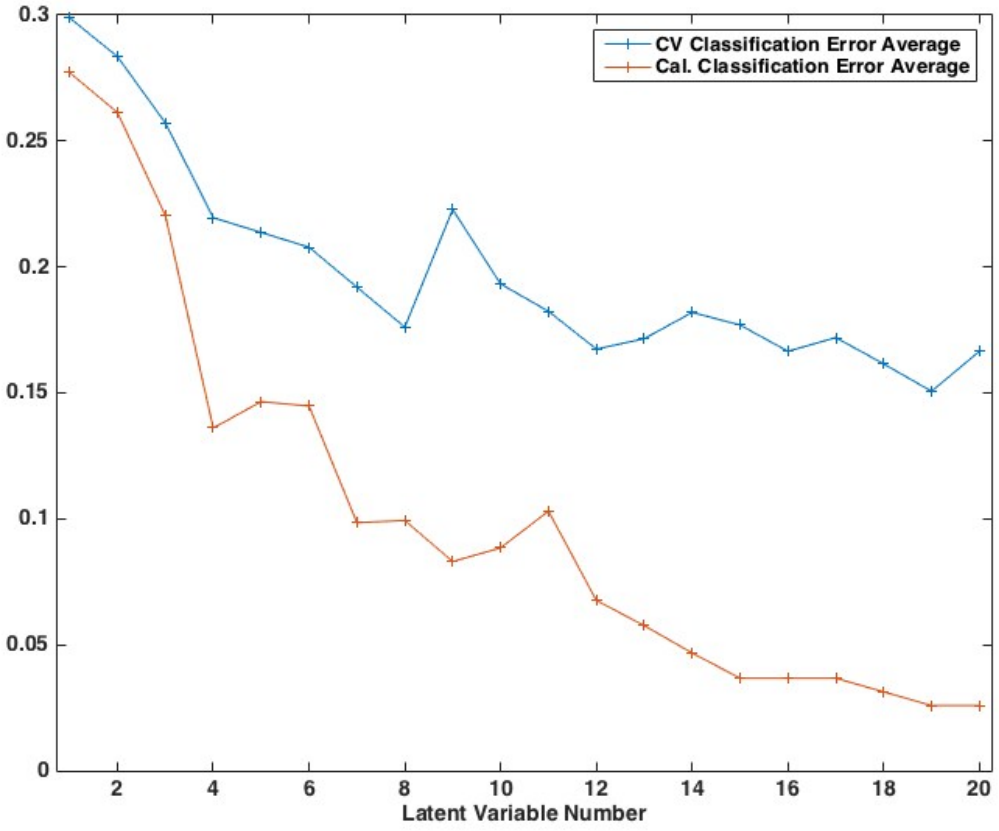


Figure S4. Error classification plots for OPLS-DA model LVs selection of 8 healthy subjects 200 urine samples ^{1}H NMR spectra before and after dictamnus ingestion.

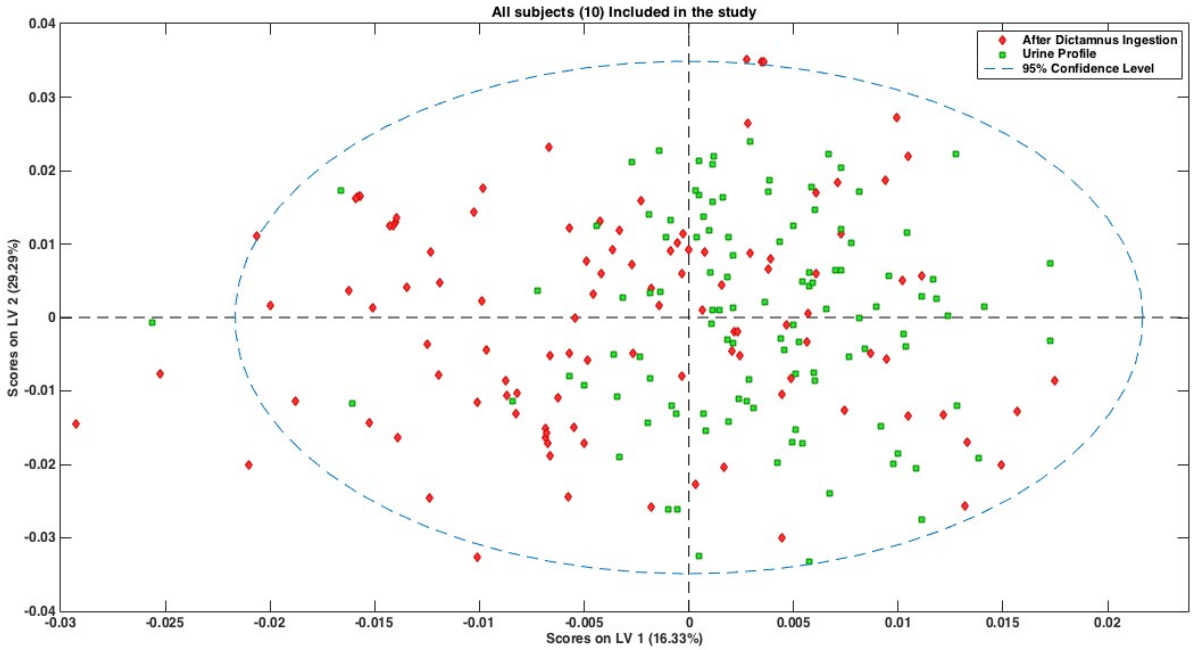


Figure S5. OPLS-DA analysis score plot of all 10 subjects (8 healthy plus 2 with chronic inflammatory bowel diseases) urine urine samples ^{1}H NMR spectra before and after dictamnus ingestion. As depicted, the two groups are not well classified (see below Table S3).

Table S3. The cross-validated values of sensitivity, specificity and accuracy of the 10 subjects' urine profile without and after dictamnus ingestion in the OPLS-DA derived model based upon their urine samples' ^{1}H-NMR spectra.

Urine samples ^{1}H-NMR 1D-noesy spectra	**After Dictamnus Ingestion**	**Urine Profile (without Dictamnus)**
Sensitivity (CV)	52.3 %	82.5 %
Specificity (CV)	82.7 %	51.8 %
Accuracy (CV)	63.6 %	74.9 %

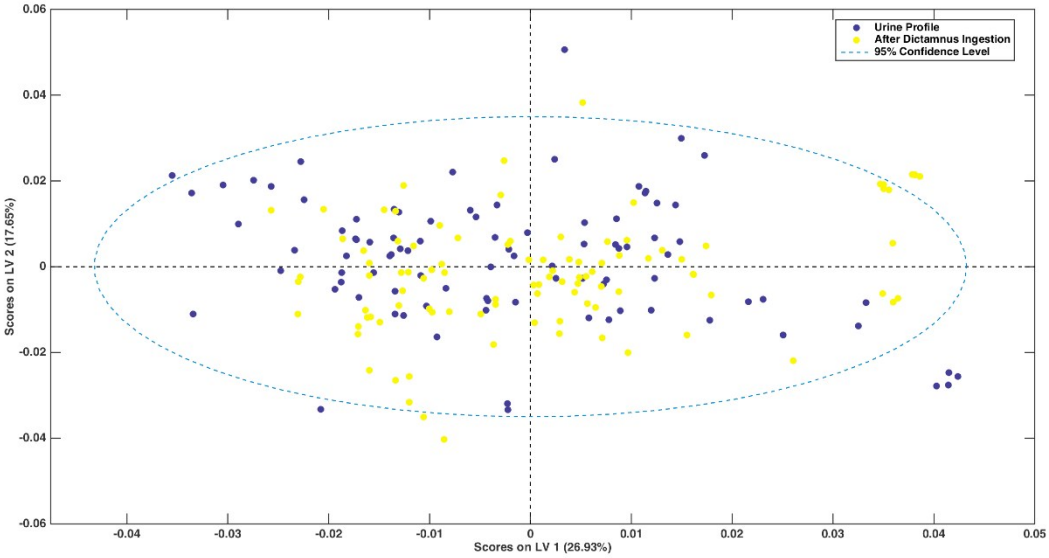


Figure S6. An alternative view of Fig. 5. All 8 healthy subjects' urine samples are colored according to before and after dictamnus tea ingestion periods. Obviously, there is not clear separation of the above mentioned periods. Even though the comparison with Fig. 1 (2 groups OPLS-DA analysis) is partially statistically unfair due to 16 groups OPLS-DA supervised analysis, this result indicates a significant contribution to each individual's metabolic profile and response to dictamnus tea ingestion.

Time series results of female subject 1 (S1) urine metabolites that exhibit significant variations after dictamnus ingestion (Figs. S7–S10).

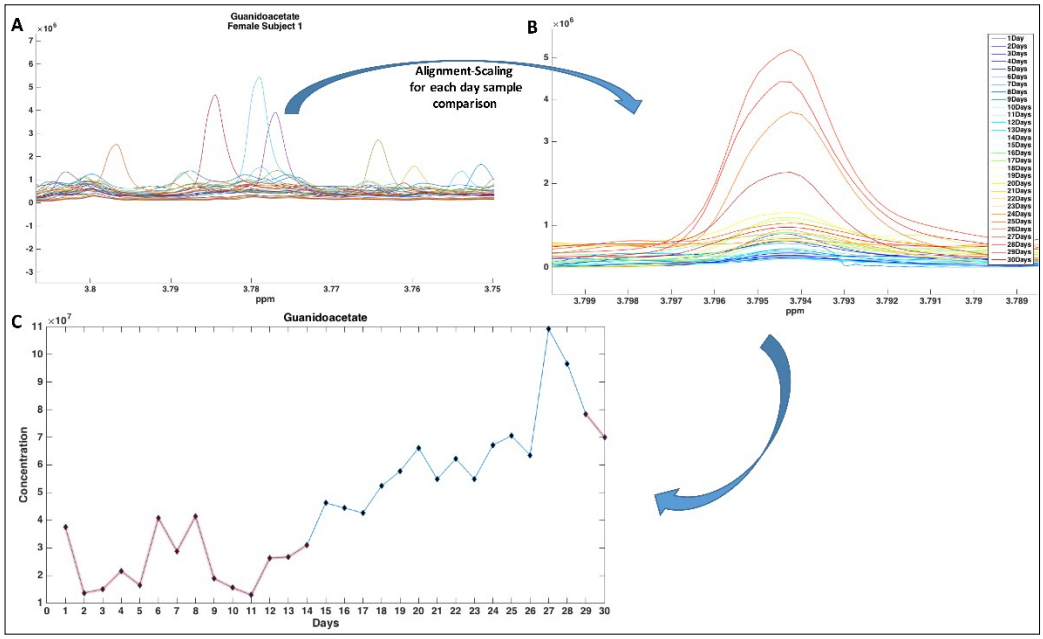


Figure S7. A) All 30 ^{1}H-NMR urine spectra collected from each day of the study period, focusing on guanidoacetate's ^{1}H-NMR spectral region. **B)** The application of our homemade algorithm for NMR peaks alignment and total area normalization of the spectra in order to achieve the lowest errors for the guanidoacetate's quantification. **C)** Guanidoacetate's concentration differences during the 30 days study period. In red are highlighted the days of no dictamnus ingestion, namely the first 2 weeks and the last 2 days.

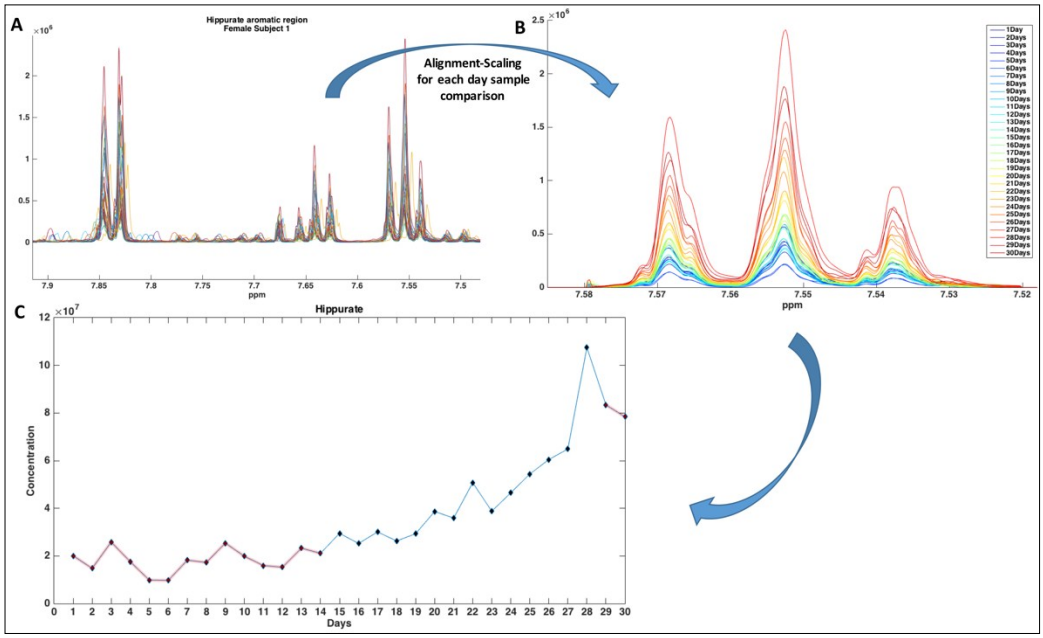


Figure S8. A) All 30 ^{1}H-NMR urine spectra collected from each day of the study period, focusing on hippurate's aromatic ^{1}H-NMR spectral region. **B)** The application of our homemade algorithm for NMR peaks alignment and total area normalization of the spectra in order to achieve the lowest errors for the hippurate's quantification. **C)** Hippurate's concentration differences during the 30 days study period. In red are highlighted the days of no dictamnus ingestion, namely the first 2 weeks and the last 2 days.

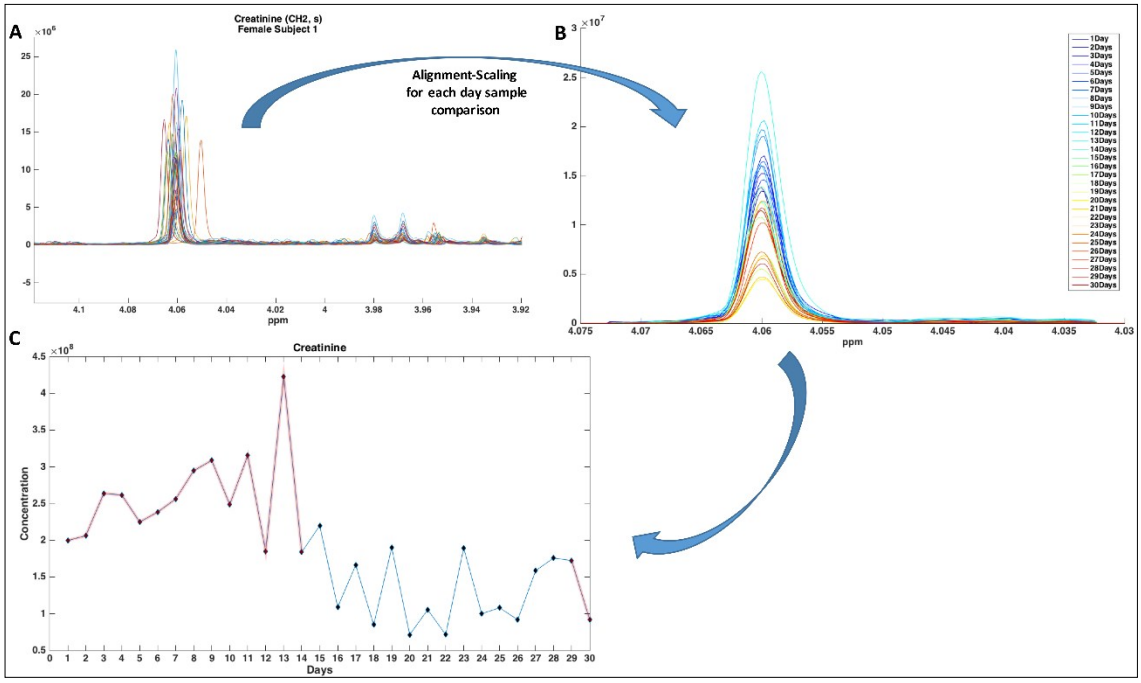


Figure S9. A) All 30 ^{1}H-NMR urine spectra collected from each day of the study period, focusing on creatinine's -CH_2 group ^{1}H-NMR peak (singlet). **B)** The application of our homemade algorithm for NMR peaks alignment and total area normalization of the spectra in order to achieve the lowest errors for the creatinine's quantification. **C)** Creatinine's concentration differences during the 30 days study period. In red are highlighted the days of no dictamnus ingestion, namely the first 2 weeks and the last 2 days.

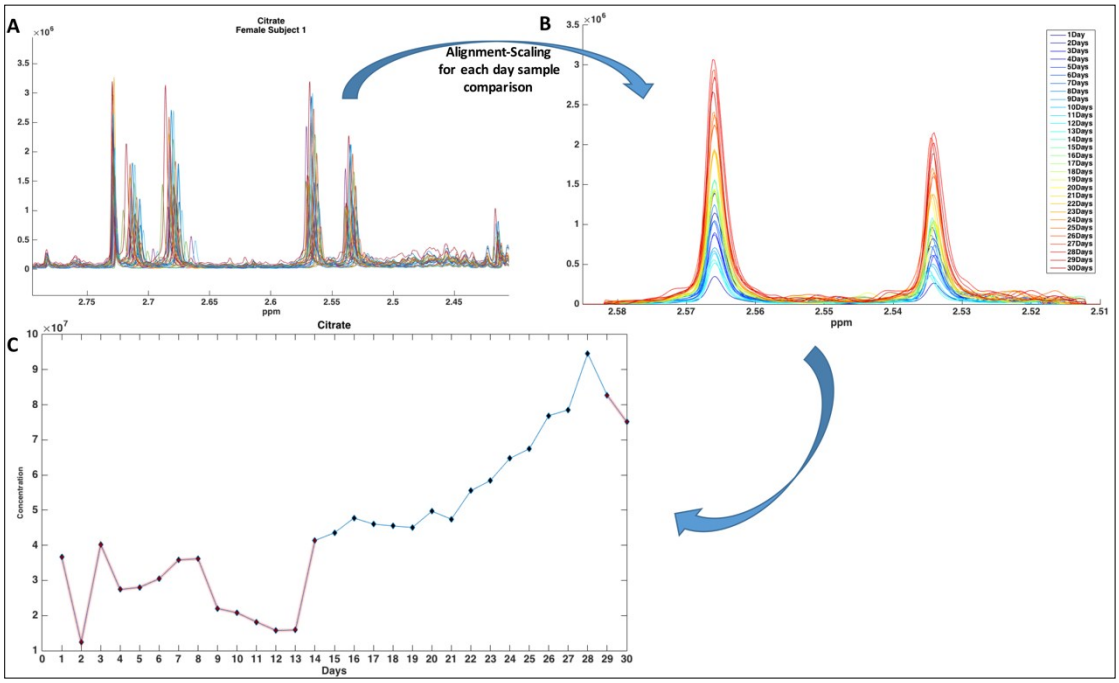


Figure S10. A) All 30 ^{1}H-NMR urine spectra collected from each day of the study period, focusing on citrate's ^{1}H-NMR spectral region. **B)** The application of our homemade algorithm for NMR peaks alignment (the one citrate's doublet) and total area normalization of the spectra in order to achieve the lowest errors for the citrate's quantification. **C)** Citrate's concentration differences during the 30 days study period. In red are highlighted the days of no dictamnus ingestion, namely the first 2 weeks and the last 2 days.

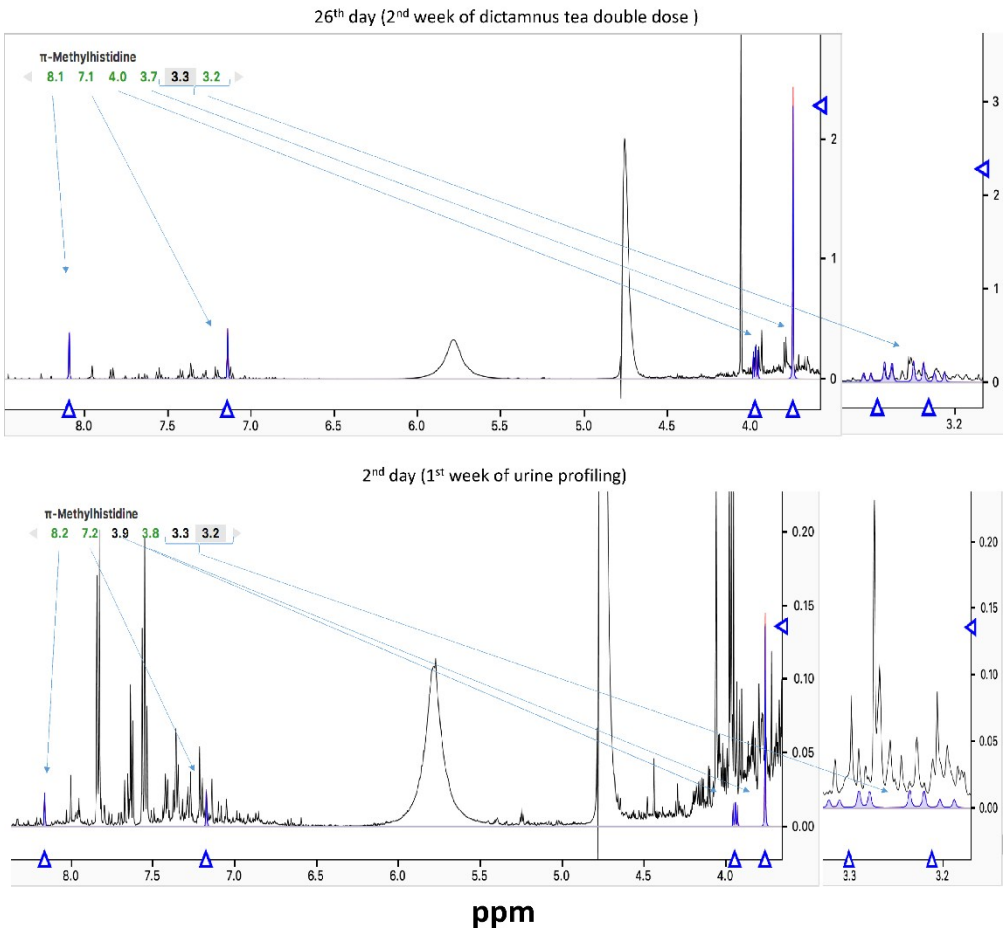


Figure S11. Chenomx platform assignment of all ^{1}H NMR spin systems of π-methyl-histidine metabolite. (Below) The female subject's 1 (S1) urine ^{1}H-NMR profile from the day 2 of the study (i.e. No dictamnus tea intake). (Upper) The female subject's 1 (S1) urine ^{1}H-NMR profile from the day 26 of the study (i.e. double dose of dictamnus tea intake).

Table S4. Concentration change of π-methyl-histidine for female subject 1 (Fig. S11) before (day 2) and after (day 26) dictamnus tea ingestion by ≈ 9 times more based upon the intensity of π-methyl-histidine' ^{1}H-NMR signal (singlet at 8.22 ppm) divided by the formate proton (singlet at 8.46 ppm) and alanine's methyl group (doublet at 1.48 ppm) intensities, since the latter exhibit no statistically significant concentration changes during the study period.

	π-methly-histidine (singlet at 8.22 ppm)		
Metabolites with no concentration changes	**Intensity ratio (Fig. S11 upper, dictamnus effect)**	**Intensity ratio (Fig. S11 down) (no dictamnus effect)**	**Concentration change**
formate (singlet at 8.46 ppm)	4.65	0.5	9.3
Alanine (doublet at 1.48 ppm)	7.73	0.92	8.4

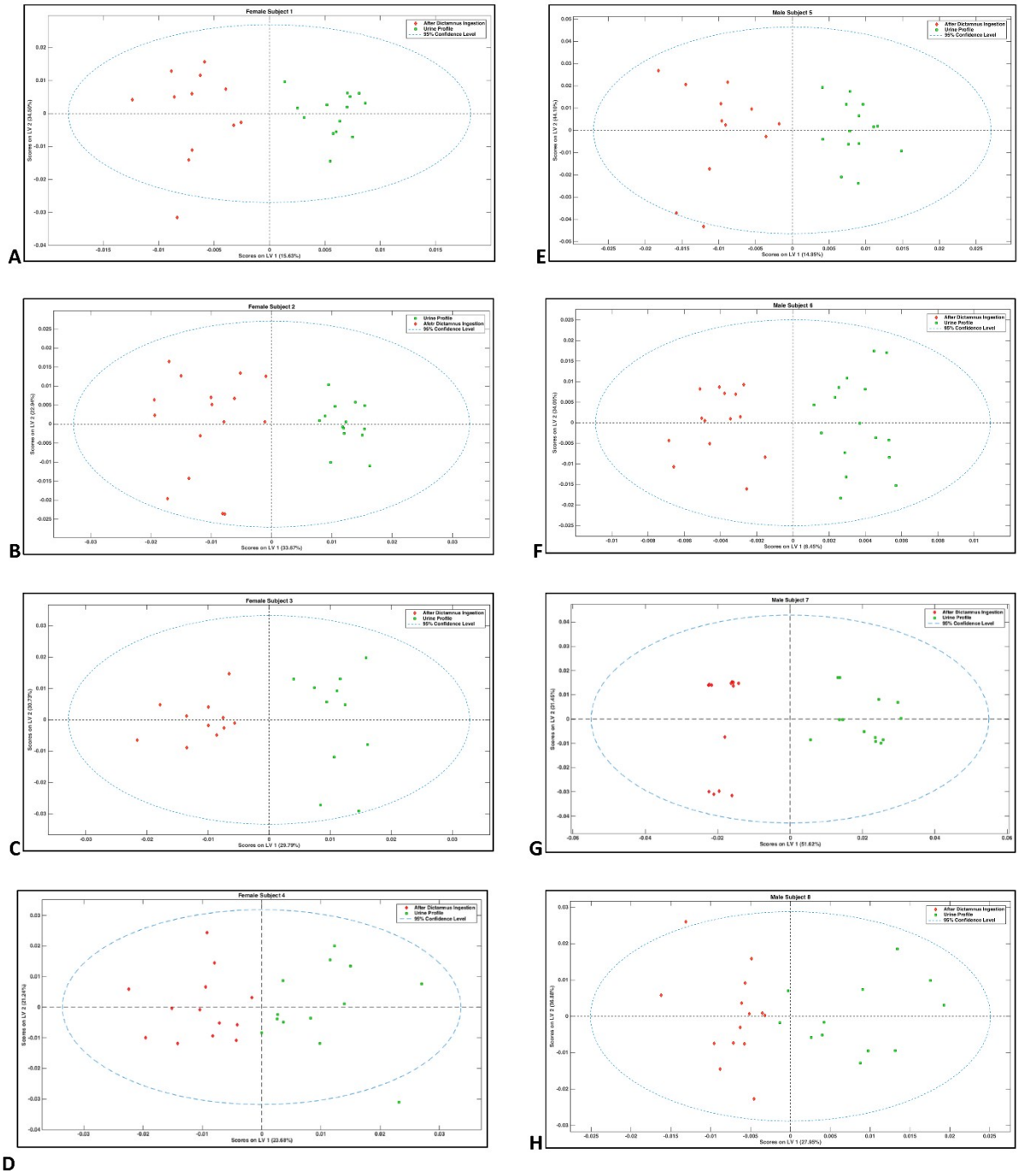


Figure S12. OPLS-DA analysis score plots from each healthy subject's (8 in total) urine samples ^{1}H-NMR spectra before and after dictamnus tea ingestion. The cases of female **A)** Subject 1 (S1), **B)** Subject 2 (S2), **C)** Subject 3 (S3), **D)** Subject 4 (S4) and of male **E)** Subject 5 (S5), **F)** Subject 6 (S6) **G)** Subject 7 (S7), **H)** Subject 8 (S8).

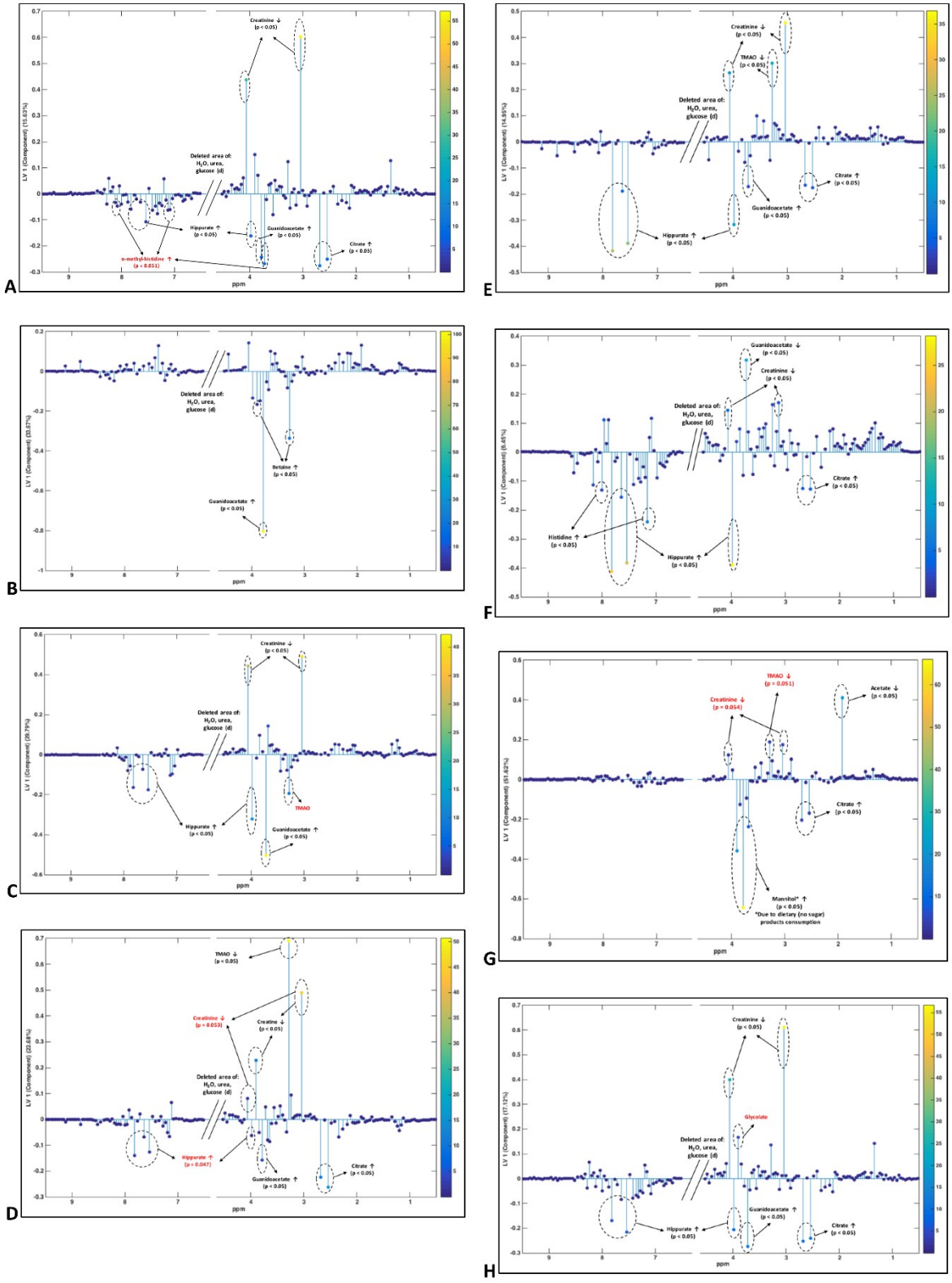


Figure S13. LV plots of the urine NMR profile without vs with dictamnus tea ingestion from the OPLS-DA analysis of each healthy subject's (8 in total) urine samples ^{1}H-NMR spectra (see Fig. S12). All assigned statistically significant ($p < 0.05$, black font), borderline significant ($p \approx 0.05$, red font) and not significant metabolites (only red font) with their concentrations trends after dictamnus ingestion (arrows) are pointed out in the LV plots. The cases of female **A)** Subject 1 (S1), **B)** Subject 2 (S2), **C)** Subject 3 (S3), **D)** Subject 4 (S4) and of male **E)** Subject 5 (S5), **F)** Subject 6 (S6) **G)** Subject 7 (S7), **H)** Subject 8 (S8).

Table S5. Cross validated (CV) sensitivity, specificity and accuracy values for each subject (8 healthy subjects) group after the OPLS-DA analysis of all subjects together before and after dictamnus ingestion (Fig. 5).

	Classification of Subjects' Urine profiles (no dictamnus) (see score plot in Fig. 5)							
Healthy Subjects	**S1**	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**	**S8**
Sensitivity (%)	83.3	93.8	91.7	72.7	83.3	85.7	100.0	90.0
Specificity (%)	93.8	92.1	97.5	91.4	85.0	89.4	98.8	86.6
Accuracy (%)	98.5	93.0	94.6	82.1	84.2	87.6	99.4	88.3
	Classification of Subjects' Urine profiles after dictamnus tea ingestion (see score plot in Fig. 5)							
Sensitivity (%)	100.0	92.9	100.0	100.0	91.7	100.0	100.0	100.0
Specificity (%)	96.6	94.2	97.8	100.0	96.6	97.6	98.9	96.5
Accuracy (%)	98.3	93.6	98.9	100.0	94.9	98.8	99.4	98.2

Time series results of male subject 10 (S10 with mild IBD) urine metabolites that exhibit significant variations after dictamnus ingestion

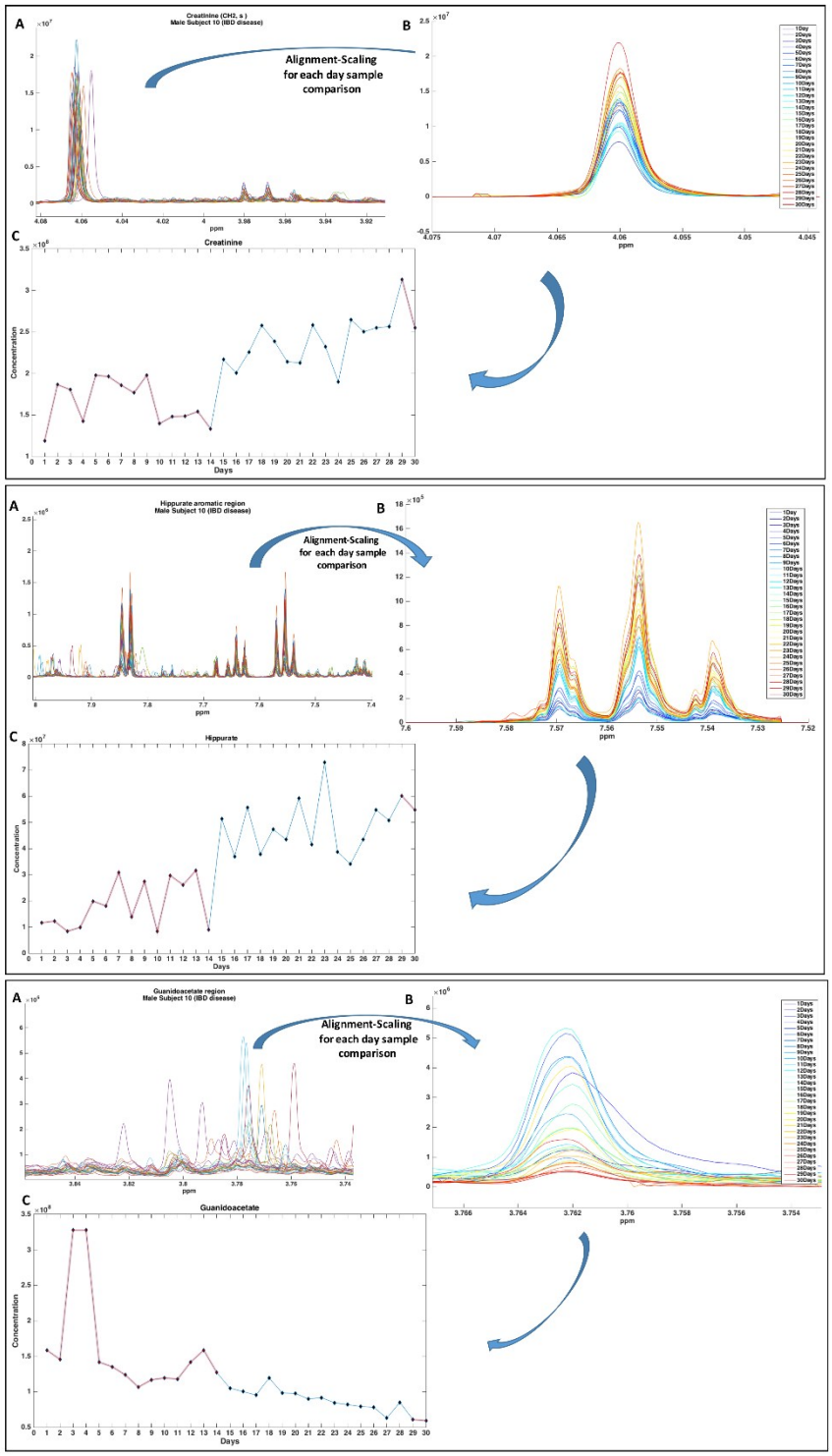


Figure S14. Time series analysis of the creatinine's (upper), hippurate's (middle) and guanidoacetate's (below) concentrations during the 30 days study period of the male subject 10 with mild IBD. **A)** All 30 ^{1}H-NMR urine spectra collected from each day of the study period, focusing on each metabolite's ^{1}H-NMR spectral region (or part of it). **B)** The application of our homemade algorithm for the best possible NMR peaks alignment and total area normalization of the spectra in order to achieve the lowest errors for the quantification of each metabolite via integration. C) Each metabolite's concentration differences during the 30 days study period. In red are highlighted the days of no dictamnus ingestion, namely the first 2 weeks and the last 2 days.

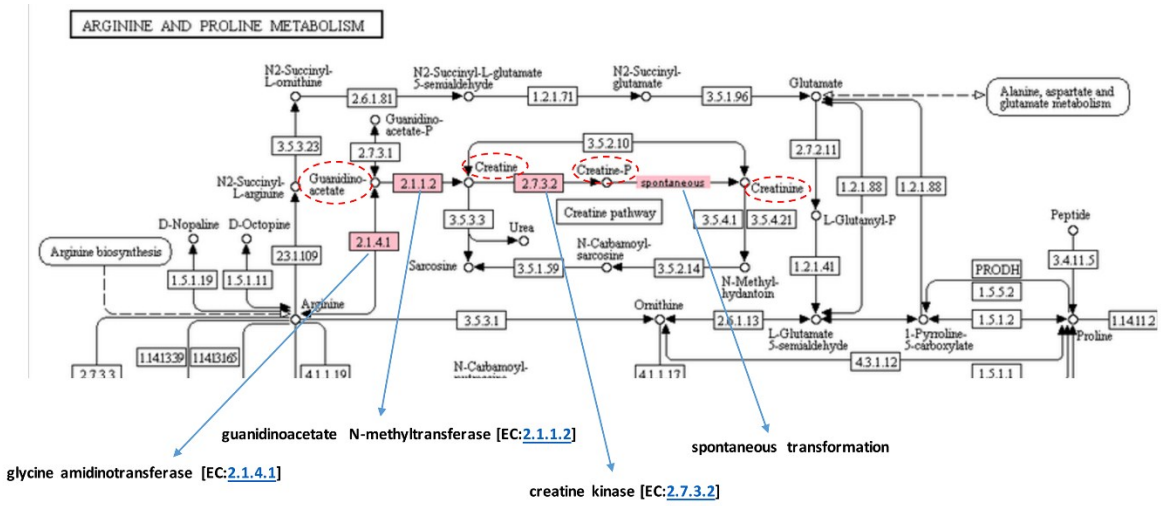


Figure S15. The metabolic pathway of arginine and proline metabolism, where the biochemical connection of guanidoacetate and creatine-creatinine formation is depicted. Data taken from KEGG pathway database [Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M., and Tanabe, M.; KEGG as a reference resource for gene and protein annotation. Nucleic Acids Res. 44, D457-D462 (2016) and Kanehisa, M. and Goto, S.; KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Res. 28, 27-30 (2000).]