

Article

Early metabolomics markers of acute low-dose exposure to uranium in rats: Supplementary Materials

Table S1. The characteristic and validation parameters of the four models calculated between the control group and the low doses of NU for different sampling times in urinary profiles. In order to identify variation in biological effects over time, four final PLS-DA models and corresponding composite scores were calculated from the urine profiles according to the treatment procedure between the control group and the low doses of NU (0.5 and 50 µg/Kg). All these models were validated in the final step after the different filtering phases and the selection of discriminating variables as summarized.

*The permutation test is considered as "very good" when the predictivity of the model decreases with the number of permutations)

	Sampling times for control/low doses of NU	PLS-DA model validation			Number of discriminant variables	Score composite validation by ROC curve
		R2Y(cum) Q2(cum)	p value	Permutation test		
Model1	24h, 48h, 5d	0.833 0.806	$1,24715 \times 10^{-25}$	Very good*	20	1
Model2	48h, 5d, 15d	0.836 0.818	$8,16391 \times 10^{-29}$	Very good*	41	1
Model3	5d, 15d, 30d	0.97 0.968	0	Very good*	33	1
Model4	15d, 30d, 90d	0.865 0.777	$1,14004 \times 10^{-16}$	Very good*	37	0.994

Table S2. New PLS-DA models calculated using the 19 common discriminant variables in at least two time-period models in urinary profiles.

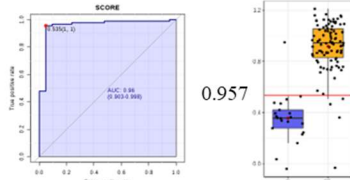
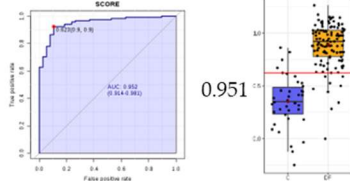
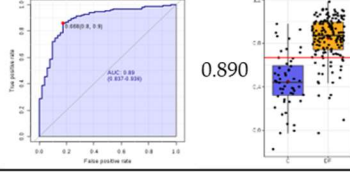
	Combining individuals from different sampling times for control/low doses of NU	PLS-DA model validation				Score composite validation by ROC curve
		R2Y(cum) Q2(cum)	Model prediction validity	p value	Permutation Test	
M' 1	From Model1 and Model2	0.735 0.653	Validated	$7,04032 \times 10^{-22}$	Very good	
M' 2	From M' 1 and Model3	0.535 0.506	Validated	$3,76288 \times 10^{-24}$	Very good	
M' 3	From M' 2 and Model4	0.395 0.37	Not validated	$1,16642 \times 10^{-20}$	—	

Table S3. The characteristic and validation parameters of the three models calculated between the control group and the low doses of NU for different sampling times in a plasma blood matrix.

	Sampling times for control/low doses of NU	PLS-DA model validation			Number of discriminant variables	Score composite validation by ROC curve
		R2Y(cum) Q2(cum)	p value	Permutation test		
Model1	48h, 5d, 15d	0.673 0.603	$6,04261 \times 10^{-16}$	Very good	30	0.934
Model2	5d, 15d, 30d	0.625 0.542	$3,36126 \times 10^{-18}$	Very good	31	0.893
Model3	15d, 30d, 90d	0.544 0.52	$6,45441 \times 10^{-18}$	Very good	28	0.944

Table S4. PLSD-DA models calculated using the 27 discriminant variables in common to at least two time models between the control group and the low doses of NU in in plasma blood matrix.

	Combining individuals from different sampling times for control/low doses of NU	PLS-DA model validation			Score composite validation by ROC curve
		R2Y(cum) Q2(cum)	p value	Permutation test	
M' 1	From Model1 and Model2	0.591 0.521	$5,94825 \times 10^{-22}$	Very good	0.878
M' 2	From M' 1 and Model3	0.322 0.27	$4,70312 \times 10^{-14}$	—	0.778

Table S5. The characteristic and validation parameters of the four models calculated between the control group and the low doses of NU for different sampling times from the annotated urinary data matrix. As observed for C18 positive ionization urinary data, sampling times were clearly differentiated with PCA applied to the annotated urinary data matrix (269 metabolites) and contaminated rats could not be differentiated from non-contaminated rats. In addition, there are no discriminant variables in common between the different time models. On the other hand, 40 discriminant variables in common to at least two models were selected, and new models were calculated based on these 40 variables. Unfortunately, none of the new reduced models calculated using these variables has been validated ($Q^2 < 0.5$) even if the ROC curve was valid. However, there are other discriminant variables characteristic of each time period, of which the most important is tryptophan.

* metabolites associated with the nicotinate-nicotinamide pathway and ** metabolites associated with the Tryptophan pathway.

	Sampling times for control/low doses of NU	PLS-DA model validation			Score composite validation by ROC curve	Discriminant variables
		R2Y(cum) Q2(cum)	p value	Permutation Test		
Model 1	24h, 48h, 5d	0.634 0.536	4.21649×10^{-18}	Very good	0.932	N,N-DIMETHYLGLYCINE D-GLUCURONOLACTONE L-GULONOLACTONE 4-QUINOLINECARBOXYLIC 7,8-DIHYDRO-L-BIOPTERIN ETHYLMALONATE 4-PYRIDOXATE 1-METHYLNICOTINAMIDE * ISO-QUINOLINE ALPHA-METHYL-DL-HISTIDINE DIHYDROCHLORIDE L-THREO-3-PHENYLSERINE(DL-3-PHENYLSERINE) O-HYDROXYHIPPURIC ACID 3-METHYLGUTARIC ACID 5-AMINOPENTANOATE 2,6-QUINOLINEDIOL (E)-3-METHYLGUTACONIC ACID NORLEUCINE N-ACETYLASPARAGINE HEXANOYL-L-CARNITINE INOSINE L-TYROSINE KYNURENIC ACID ** PROLINE
Model 2	48h, 5d, 15d	0.524 0.502	3.99664×10^{-13}	Very good	0.939	5-AMINOPENTANOATE INDOLE-3-ACETATE ** L-THREO-3-PHENYLSERINE NORLEUCINE L-GULONOLACTONE HEXANOYL-L-CARNITINE 2-ISOPROPYLMALIC ALPHA-METHYL-DL-HISTIDINE 3,4-DIHYDROXY-L-PHENYLALANINE SEROTONIN ** HYDROXYPHENYLACTATE DL-BENZYL SUCCINIC N-ACETYLMETHIONINE (E)-3-METHYLGUTACONIC ACID TARTRATE 4-METHYL-2-OXOVALERIC ACID 5-HYDROXYTRYPTOPHAN ** AMINOADIPATE N-ACETYL-L-TYROSINE 3-(2-HYDROXYPHENYL)PROPANOATE N-ALPHA-ACETYL-LYSINE
Model 3	5d, 15d, 30d	0.635 0.567	1.39002×10^{-20}	Very good	0.934	N,N-DIMETHYLGLYCINE 5-METHYLCYTOSINE D-GLUCURONOLACTONE 3-(2-HYDROXYPHENYL)PROPANOATE SACCHARATE TRYPTOPHAN ** DL-BENZYL SUCCINIC CYTOSINE SORBITOL ADIPATE L-HISTIDINE THIAMINE PENTOSE GLUTARATE OXOGLUTARATE ** MALATE BETA-ALANINE QUINOLINATE * / ** HYPOTAURINE 3-HYDROXYBENZOATE L-LYSINE URATE HYDROXYPHENYLACTATE 1-AMINOCYCLOHEXANECARBOXYLIC
Model 4	15d, 30d, 90d	0.64 0.551	2.13671×10^{-19}	Very good	0.959	BETA-ALANINE DL-BENZYL SUCCINIC L-THREO-3-PHENYLSERINE 3-HYDROXYBENZOATE D-GLUCURONOLACTONE CYTOSINE 3-METHYL-2-OXINDOLE DELTA PROLINE L-GULONOLACTONE 3-(2-HYDROXYPHENYL)PROPANOATE ALPHA-METHYL-DL-HISTIDINE L-PHENYLALANINE L-SERINE HYPOTAURINE L-GLUTAMIC ACID *** L-HISTIDINE N-FORMYL-L-METHIONINE L-LYSINE (E)-3-METHYLGUTACONIC ACID INDOLE-3-ACETATE ** TRYPTOPHAN ** SORBITOL BETAINE 3-METHYL-2-OXOVALERATE GALACTURONATE 5-AMINOPENTANOATE

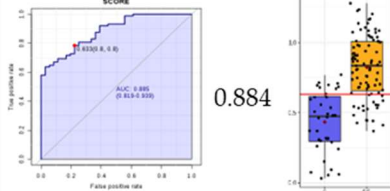
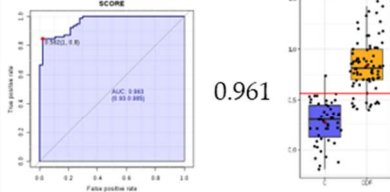
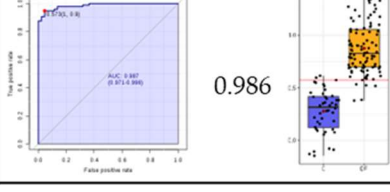
Table S6. The characteristic and validation parameters of the three models calculated between the control group and the low doses of NU for different sampling time in an annotated plasma matrix. Similar PCA results to C18 positive ionization data were observed and the same data procedure was applied for the annotated plasma matrix where three validated models were calculated. Six discriminant metabolites were in common between all the three models with overlapping time windows and therefore can be considered as time-independent biomarkers. These metabolites were putatively identified as cholate (level 1 according to MSI), nicotinamide mononucleotide (level 3), dimeric cysteine (level 1), 5b-cholanic acid-3a-ol-12-one (level 3), 5b-cholanic acid-3a.12a-

diol-7 (level 3) and citrulline (level 1). In addition, the most interesting result from the combination of the different columns and ionization modes was the detection of new potential biomarkers such as tryptophan and nicotinamide mononucleotide, which are related to the N-methylnicotinamide pathway.

* metabolites associated with the nicotinate-nicotinamide pathway and ** metabolites associated with the Tryptophan pathway.

	Sampling times for control/low doses of NU	PLS-DA model validation			Score composite validation by ROC curve	Discriminant variables
		R2Y(cum) Q2(cum)	p value	Permutation n Test		
Model1	48h, 5d, 15d	0.61 0.532	1.24232×10^{-15}	Very good	0.99	L-PHENYLALANINE NICOTINAMIDE MONONUCLEOTIDE CYSTEINE DIMERE P-OCTOPAMINE 2,6-DIHYDROXYBENZOIC ACID L-HISTIDINE CHOLATE L-TYROSINE GLY-PRO 5B-CHOLANIC ACID-3A-OL-12-ONE L-ASPARAGINE L-SERINE PROLINE CITRULLINE 2,2-DIMETHYLSUCCINIC ACID 5B-CHOLANIC ACID-3A.12A-DIOL-7-ONE D-ORNITHINE
Model2	5d, 15d, 30d	0.623 0.549	1.45847×10^{-17}	Very good	0.923	5B-CHOLANIC ACID-3A-OL-12-ONE CHOLATE 5B-CHOLANIC ACID-3A.12A-DIOL-7-ONE 5B-CHOLANIC ACID-3A.6A.7A-TRIOL 5B-CHOLANIC ACID-3A.12A-DIOL-7-ONE NICOTINAMIDE MONONUCLEOTIDE* DEOXYCHOLATE 5B-CHOLANIC ACID-3A-OL-12-ONE 5B-CHOLANIC ACID-3A-OL-7.12-DIONE 2-OXOADIPATE GLYCOCHOLATE PIPECOLATE TRIMETHYLAMINEN-OXIDE 2-AMINOISOBUTYRATE CITRULLINE L-TYROSINE 5-METHYLCYTOSINE OXOGLUTARATE** CYSTEINE DIMERE MEVALONOLACTONE HYDRATE P-OCTOPAMINE (S)-3-HYDROXYBUTYRIC ACID-H- ET (S)-3-HYDROXYBUTYRIC ACID 16-0 LYSOPI
Model3	15d, 30d, 90d	0.596 0.518	9.70734×10^{-16}	Very good	0.95	NICOTINAMIDE MONONUCLEOTIDE 5B-CHOLANIC ACID-3A-OL-12-ONE 5B-CHOLANIC ACID-3A.12A-DIOL-7-ONE CHOLATE 18-0 LYSOPC CITRULLINE 5B-CHOLANIC ACID-3A.6A.7A-TRIOL DEOXYCHOLATE 16-0 LYSOPC ALLANTOIN 5-HYDROXYINDOLE-3-ACETIC-ACID** 2,6-QUINOLINEDIOL TRYPTOPHAN** C16 PLASMLPC CYSTEINE DIMERE C18 LYSOPAF 18-0 LYSOPE D-ORNITHINE

Table S7. The characteristic and validation parameters of the three models calculated between the control group and the low doses of NU for different sampling time in annotated urinary and plasma data matrices.

	Sampling times for control/low doses of NU	PLS-DA model validation			Number of discriminant variables	Score composite validation by ROC curve
		R2Y(cum) Q2(cum)	p value	Permutation test		
Model 1	48h, 5d, 15d	0.714 0.597	4.02985×10^{-21}	Very good	52	 0.884
Model 2	5d, 15d, 30d	0.791 0.64	6.56191×10^{-26}	Very good	49	 0.961
Model 3	15d, 30d, 90d	0.678 0.57	6.82626×10^{-24}	Very good	53	 0.986

Tables S8 a, b, c, d. The discriminant variables detected in common to at least two models calculated between the control group and the low doses of NU for different sampling times for annotated urinary and plasma data.* Model calculated between control samples and low dose NU for different sampling times between 48h, 5d and 15d; 5d, 15d and 30d; 15d, 30d and 90d.

		48 hours to 15 days				5 days to 30 days				15 days to 90 days			
a. 6 common discriminant variables between 48 hours to 15 days, 5 days to 30 days and 15 days to 90 days		Fold Change	p-value	FDR	Box plot	Fold Change	p-value	FDR	Box plot	Fold Change	p-value	FDR	Box plot
Urine_CP_M137T39	3-METHYLNICOTINAMIDE-H+	4.5475	5.93×10^{-11}	3.4×10^{-11}		3.0335	1.30×10^{-12}	2.30×10^{-12}		3.3398	1.99×10^{-11}	5.27×10^{-11}	
Urine_CP_M190T38	BETA-ALANINE-H+	0.35182	1.25×10^{-10}	4.76×10^{-11}		0.41475	4.53×10^{-11}	2.78×10^{-11}		0.22336	1.94×10^{-11}	5.14×10^{-11}	
Urine_CN_M121T43	D-GLUCURONOLACTONE-HCOOH-H-	0.42387	3.11×10^{-11}	3.59×10^{-11}		0.16749	3.98×10^{-11}	4.57×10^{-11}		0.35612	1.74×10^{-11}	4.57×10^{-11}	
Urine_CP_M104T39_1	N,N-dimethylglycine-H+	0.34799	1.27×10^{-11}	2.64×10^{-11}		0.18562	1.40×10^{-11}	3.80×10^{-11}		0.2196	2.21×10^{-11}	3.27×10^{-11}	
Urine_CN_M120T40	SACCHARATE-H-	0.1107	2.26×10^{-11}	3.96×10^{-11}		0.40602	5.05×10^{-11}	3.96×10^{-11}		0.42062	4.11×10^{-11}	2.15×10^{-11}	
Plasma_CP_M166T208 and CP_M120T208 and CP_M149T208 or Urine_CP_M166T209 and CP_M120T209 and CP_M149T209	L-Phenylalanine-H+ and L-Phenylalanine-HCOOH-H+ and L-Phenylalanine-NH3-H+	0.55505	8.19×10^{-11}	1.29×10^{-11}		0.37348	2.07×10^{-11}	4.39×10^{-11}		0.23952	1.95×10^{-11}	1.44×10^{-11}	
b. 5 common discriminant variables between 48 hours to 15 days and 5 days to 30 days		Fold Change	p-value	FDR	Box plot	Fold Change	p-value	FDR	Box plot	Fold Change	p-value	FDR	Box plot
Urine_CN_M145T258 and CN_M101T359	ADIPATE-H- and ADIPATE-H-CO2	2.6988	4.24×10^{-11}	1.05×10^{-11}		2.1539	6.23×10^{-11}	1.63×10^{-11}					
Urine_CN_M133T46 and CN_M115T46	MALATE-H- and MALATE-H2O-H-	2.623	4.13×10^{-11}	1.05×10^{-11}		2.2175	2.57×10^{-11}	4.80×10^{-11}					
Plasma_HP_M424T124	Sb-cholanic acid-3a-12a-DIOL-7-ONE-NH4+	0.4261	2.20×10^{-11}	2.66×10^{-11}		0.34482	1.87×10^{-11}	4.25×10^{-11}					
Plasma_HP_M355T137	Sb-cholanic acid-3a-OL-12-ONE-H2O-H+	0.51625	2.70×10^{-11}	3.19×10^{-11}		0.42487	1.23×10^{-11}	3.02×10^{-11}					
Plasma_CN_M475T491 and CN_M443T491 and CN_M467T491 and CN_M455T491	CHOLATE-NaHCO2-H- and CHOLATE-Cl- and CHOLATE-H- and CHOLATE-HCOOH-H-	0.46858	1.07×10^{-11}	1.43×10^{-11}		0.46869	4.25×10^{-11}	3.77×10^{-11}					
c. 3 common discriminant variables between 48 hours to 15 days and 15 days to 90 days		Fold Change	p-value	FDR	Box plot	Fold Change	p-value	FDR	Box plot	Fold Change	p-value	FDR	Box plot
Urine_CP_M115T53	5-AMINOPENTANOATE-H+	0.28797	5.55×10^{-11}	5.69×10^{-11}						0.25579	9.29×10^{-11}	1.27×10^{-11}	
Urine_CN_M159T299 and CN_M115T300	6-CARBONYHEXANOATE-H- and 6-CARBONYHEXANOATE-H-CO2	2.2581	1.63×10^{-11}	3.26×10^{-11}						1.7233	8.38×10^{-11}	1.27×10^{-11}	
Urine_CP_M144T297 or Plasma_CP_M144T264	TRYPTAMINE-NH3-H+	1.4268	2.02×10^{-11}	2.50×10^{-11}						0.45513	1.30×10^{-11}	1.44×10^{-11}	
d. 22 common discriminant variables between 5 days to 30 days and 15 days to 90 days		Fold Change	p-value	FDR	Box plot	Fold Change	p-value	FDR	Box plot	Fold Change	p-value	FDR	Box plot
Urine_HP_M196T134	2-HYDROXYPYRIDINE-H+					2.5202	2.15×10^{-11}	4.39×10^{-11}		2.4274	2.23×10^{-11}	5.27×10^{-11}	
Urine_HP_M165T115 and HP_M147T115	3-(2-HYDROXYPHENYL)PROPANOATE-H- and 3-(2-HYDROXYPHENYL)PROPANOATE-H2O-H-					3.2462	5.75×10^{-11}	5.64×10^{-11}		2.3569	2.29×10^{-11}	5.27×10^{-11}	
Urine_CN_M153T294	3-HYDROXYBENZOATE-HCOOH-H-					0.16546	4.85×10^{-11}	1.49×10^{-11}		0.15749	1.54×10^{-11}	1.43×10^{-11}	
Urine_CP_M134T391	5-Hydroxyindole-H+					3.4999	1.05×10^{-11}	2.94×10^{-11}		2.6697	9.91×10^{-11}	1.27×10^{-11}	
Urine_CP_M126T46_2	5-METHYLCYTOSINE-H+					3.0805	1.30×10^{-11}	3.17×10^{-11}		2.2542	5.65×10^{-11}	2.71×10^{-11}	
Urine_CP_M115T40	BETANINE-H+					0.17949	2.65×10^{-11}	4.80×10^{-11}		0.25579	1.15×10^{-11}	3.80×10^{-11}	
Urine_CP_M112T40	CYTOSINE-H+					3.4712	2.63×10^{-11}	4.29×10^{-11}		2.6671	3.96×10^{-11}	7.00×10^{-11}	
Urine_CP_M120T375	di-benzylsuccinic_acid-H+					0.17403	1.23×10^{-11}	9.80×10^{-11}		0.29352	1.02×10^{-11}	1.35×10^{-11}	
Urine_CN_M121T39_2 and CN_M122T36 and CN_M151T38	SORBITOL-Cl- and SORBITOL-HCOOH-H- and SORBITOL-H-					3.9214	7.39×10^{-11}	3.17×10^{-11}		2.457	6.45×10^{-11}	2.71×10^{-11}	
Urine_HP_M110T536	HYPOTAUDINE-H+					0.26261	2.89×10^{-11}	1.10×10^{-11}		0.27312	2.63×10^{-11}	3.80×10^{-11}	
Urine_CP_M176T374 and CP_M130T374	INDOLE-3-ACETATE-H+ and INDOLE-3-ACETATE-HCOOH-H+					0.2941	1.21×10^{-11}	3.02×10^{-11}		0.20942	1.62×10^{-11}	1.43×10^{-11}	
Urine_CN_M155T40	Pentose_1					2.4023	2.91×10^{-11}	1.10×10^{-11}		1.6262	8.76×10^{-11}	1.27×10^{-11}	
Urine_CP_M166T209 and CP_M120T209 and CP_M149T209	L-Phenylalanine-H+ and L-Phenylalanine-HCOOH-H+ and L-Phenylalanine-NH3-H+					0.37348	2.07×10^{-11}	4.39×10^{-11}		0.23952	1.95×10^{-11}	1.44×10^{-11}	
Urine_CP_M152T92_1	L-Threo-3-Phenylserine-DL-3-Phenylserine-H+					0.35991	7.66×10^{-11}	3.93×10^{-11}		0.14457	2.45×10^{-11}	1.44×10^{-11}	
Urine_CN_M305T40	N-Acetylneuraminic_acid-H					0.3367	4.65×10^{-11}	3.91×10^{-11}		0.27138	7.67×10^{-11}	2.71×10^{-11}	
Urine_CN_M120T343	N-ACETYLPHENYLALANINE-H-					3.6120	5.93×10^{-11}	7.52×10^{-11}		3.3627	1.03×10^{-11}	1.27×10^{-11}	
Urine_CP_M124T361	N-ACETYLTRYPHOPHAN-H+					2.0207	5.64×10^{-11}	7.52×10^{-11}		1.6565	1.04×10^{-11}	1.27×10^{-11}	
Urine_CP_M116T42	PROLINE-H+					0.31267	4.12×10^{-11}	1.35×10^{-11}		0.16027	9.06×10^{-11}	4.80×10^{-11}	
Urine_CN_M166T56 and CN_M122T56	QUINOLINATE-H- and QUINOLINATE-H-CO2					0.42288	5.63×10^{-11}	7.52×10^{-11}		0.41909	2.80×10^{-11}	3.90×10^{-11}	
Urine_HP_M120T563 and HP_M155T654	TRYPTOPHAN-H+ and TRYPTOPHAN-NH3-H+					0.22264	5.36×10^{-11}	2.92×10^{-11}		0.4117	9.62×10^{-11}	1.27×10^{-11}	
Plasma_CN_M157T39	ALLANTOIN-H-					1.9932	5.95×10^{-11}	7.52×10^{-11}		2.5354	1.22×10^{-11}	3.80×10^{-11}	
Plasma_CP_M130T52 and CP_M54T51	PIPECOLATE-H+ and PIPECOLATE-HCOOH-H+					2.0949	3.73×10^{-11}	1.30×10^{-11}		1.5738	7.56×10^{-11}	1.27×10^{-11}	

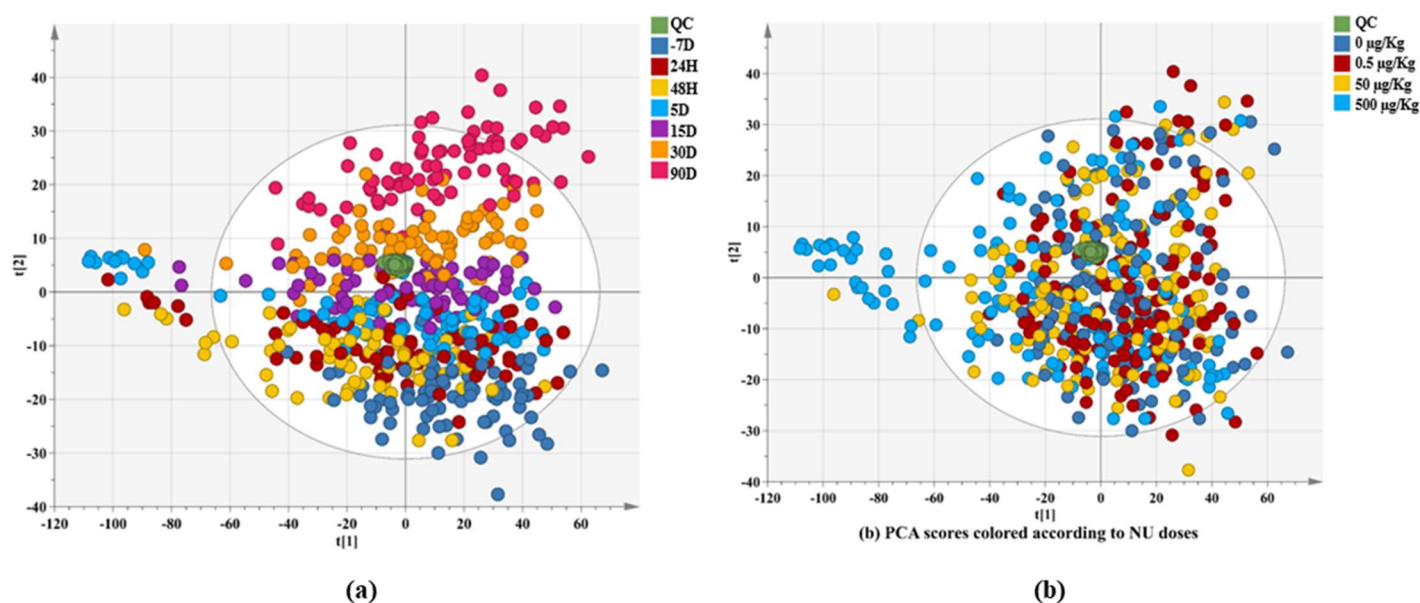


Figure S1. PCA performed on all urine sampling times and NU doses. PCA is a descriptive analysis method used to visualize the distribution of samples in a low dimensional space and demonstrate a possible structuring in the data without any information on the nature of samples. **Figure S1a,b** are the same but just colored differently according to (a) sampling time regardless of uranium dose groups and (b) to NU doses regardless of the time collection point without considering this information during the calculation of PCA. In (a), the sampling time effect is visible on the second PCA component.

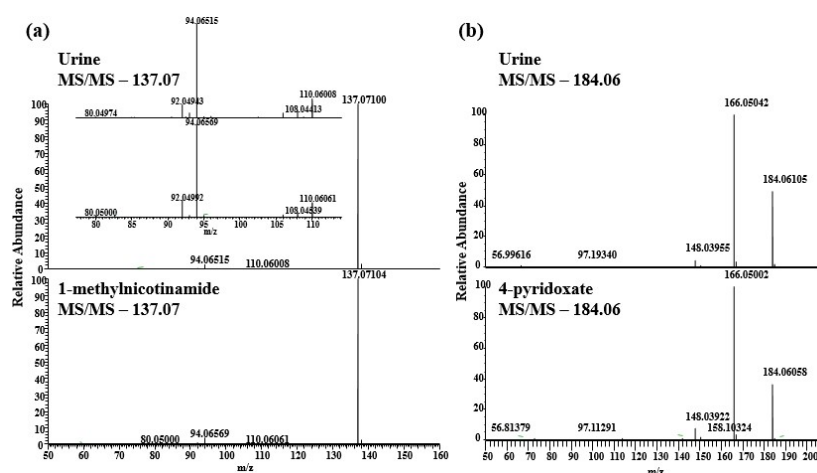


Figure S2. MS/MS spectra of the precursor $[M + H]^+$ ions at (a) m/z 137 putatively identified as 1-methylnicotinamide and (b) m/z 184 putatively identified as 4-pyridoxate from the authentic standard and from the pool urine sample.

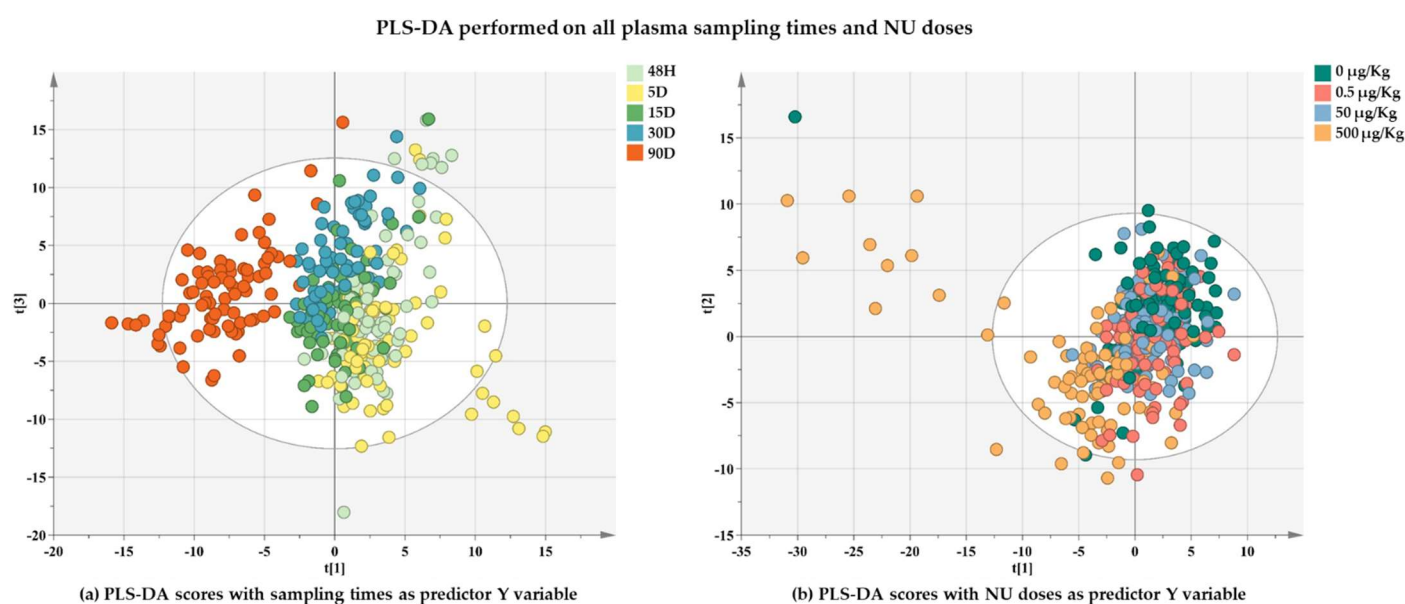


Figure S3. PLS-DA performed on all plasma sampling times and NU dose samples. **(a)** Score plot with sampling times as the predictor Y variable. **(b)** Score plot with NU doses as the predictor Y variable. PCA is a descriptive analysis method used to visualize the distribution of samples in a low dimensional space and demonstrate a possible structuring in the data without any information on the nature of samples. **Figure S2a,b** are the same but just colored differently according to **(a)** sampling time regardless of uranium dose groups and **(b)** to NU doses regardless of the time collection point without considering this information during the calculation of PCA.

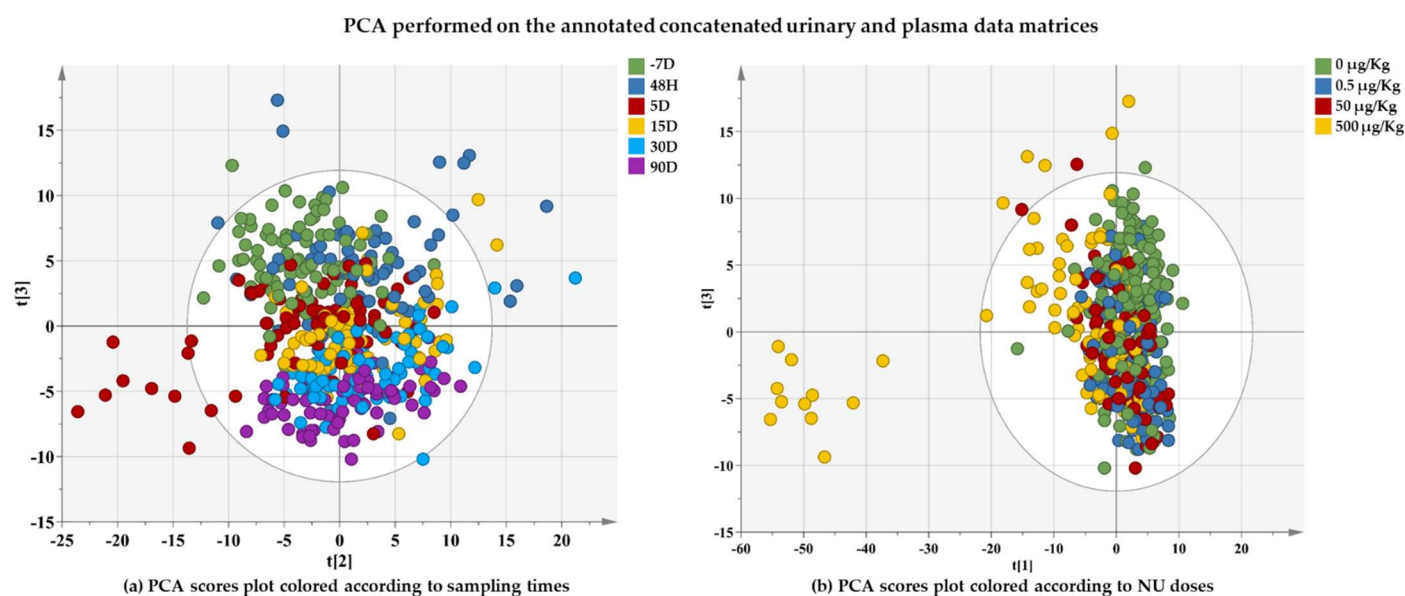


Figure S4. PCA performed on all sampling times and NU dose samples on the annotated concatenated urinary and plasma data matrices **(a)** Score plot colored according to sampling time. **(b)** Score plot colored according to NU dose. **Figure S3a,b** are the same but just colored differently according to **(a)** sampling time regardless of uranium dose groups and **(b)** to NU doses regardless of the time collection point without considering this information during the calculation of PCA.