SUPPLEMENTARY DATA

A two-way interaction between methotrexate and the gut microbiota of male Sprague Dawley rats

Marine P. M. Letertre[1,], Nyasha Munjoma[2], Kate Wolfer[1], Alexandros Pechlivanis[1,3,4], Julie A. K. McDonald[1], Rhiannon N. Hardwick[5], Nathan J. Cherrington[5], Muireann Coen[1,6], Jeremy K. Nicholson[7], Lesley Hoyles[1,8], Jonathan R. Swann[1], Ian D. Wilson[1].*

1. Imperial College, London, SW7 2AZ, UK

2. Waters Corporation, Wilmslow, SK9 4AX, UK.

3. Center for Interdisciplinary Research of the Aristotle University of Thessaloniki (KEDEK), 57001 Thessaloniki, Greece.

4. Department of Chemistry, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece

5. Department of Pharmacology and Toxicology, College of Pharmacy, University of Arizona, Tuscon, Arizona, USA.

6. Oncology Safety, Clinical Pharmacology & Safety Sciences, R&D, Astra Zeneca, Cambridge.

7. Australian National Phenome Centre, Health Futures Institute, Murdoch University, Murdoch, WA6150, Australia.

8. Department of Biosciences, Nottingham Trent University, Nottingham, NG11 8NS, UK.

Corresponding authors*E-mail: marine.letertre@biogenouest.org *E-mail: i.wilson@imperial.ac.uk

Contents

Figure S3. CA-PLS-R (covariate-adjusted projection to latent structures regression) analysis performed on the urine samples over the time points 0, 12, 24 and 48h for each of the MTX dose- groups.

Figure S4. Skyline significance of the CA-PLS-Regression models, performed on the urine samples for each of the MTX dose groups (0, 10, 40 and 100 mg/kg), representing the number of features having a q-value less than 0.05 and which significantly increased (red colour) or decreased (blue colour) over time (0, 12, 24 and 48h).

Figure S5. PCA score plots of the faecal metabolic profiles plotted according the dose of MTX administered.

Figure S6. CA-PLS-R (covariate-adjusted projection to latent structures regression) analysis performed over time (0, 12 and 48h) on the faecal samples of each of the MTX dose groups.

Figure S7. Skyline significance of the CA-PLS-Regression models, performed on faecal samples for each of the MTX dose groups (0, 10, 40 and 100 mg/kg), representing the number of features having a q-value lower than 0.05 and which significantly increased (red colour) or decreased (blue colour) over time (0, 12 and 48h).

Figure S8. β diversity represented using PCoA of Bray-Curtis distance, displaying the similarity between the samples.

Figure S9. Correlogram integrating the putatively annotated endogenous compounds excreted in urine and feces samples of the animals pre-MTX treatment with MTX concentration measured in urine at 6-12 h and in feces at 36-48 h, as well as the quantity of DAMPA excreted in feces samples at 36-48 h.

Table S1. Number of reads per samples after 16S rRNA gene sequencing pre-processing and prior importation into phyloseq for data investigation.

Table S2. Putative annotations for metabolites detected in the urine samples.

Table S3. Putative annotations for metabolites detected in the faecal samples.

Table S4. α diversity values representing the microbial richness of each of the faecal samples of the animals according the MTX dose groups over time.

R script for Feature Removal:

The following script was used to remove LC-MS-detected features that were found by multivariate statistical analysis to be significantly changed overtime in the control groups, and therefore unrelated to MTX exposure, from the significant features changed in the MTX-dosed groups.

```
setwd("…")

X <- read.csv("Matrice.csv", header = TRUE)

X.2 <- X[!X$m.z %in% X$m.z[X$Group == "Control"], ]

write.csv(X.2, file= "Matrice_minus_control.csv")
```

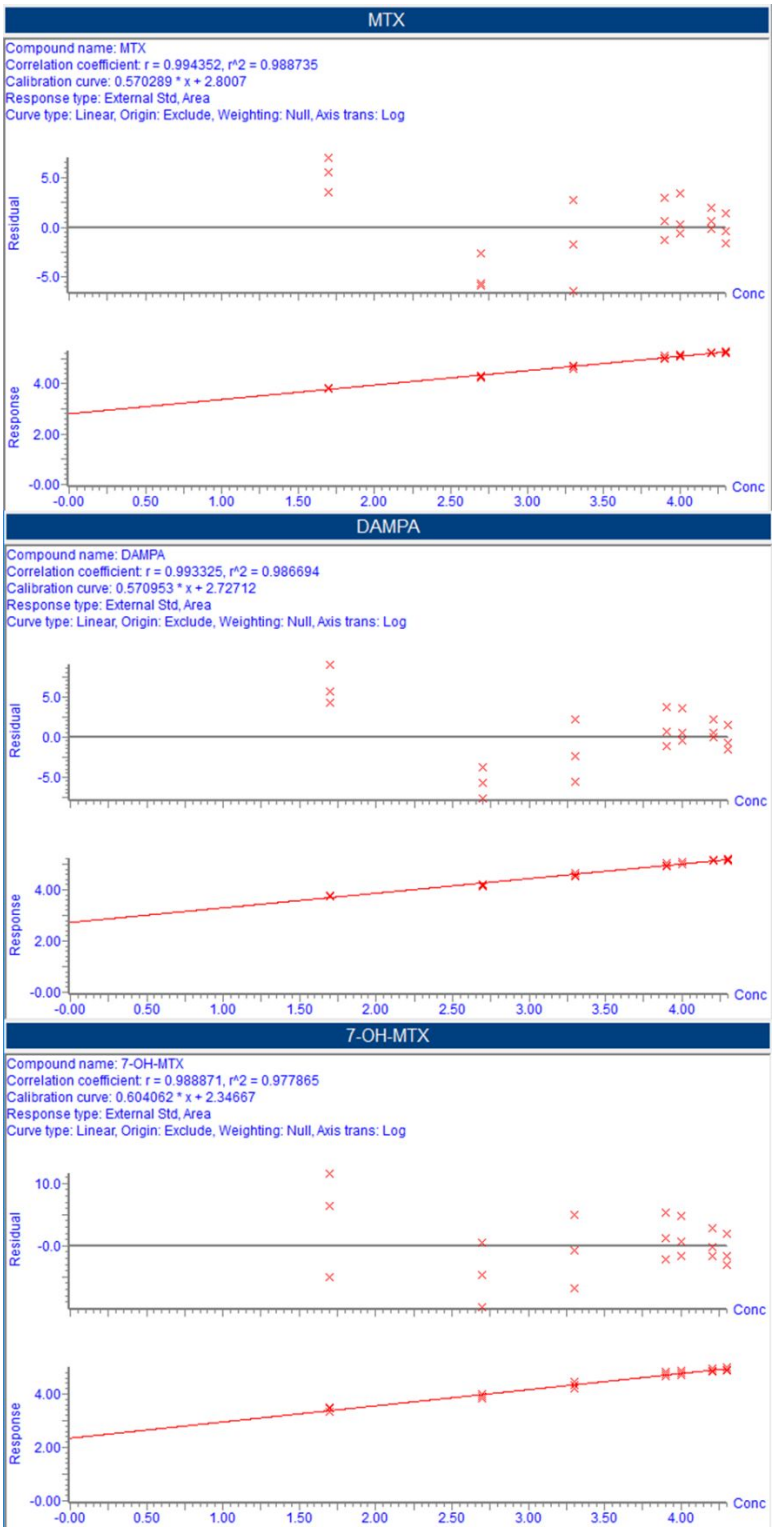

Figure S1. Screenshot of TargetLynx V4.1 showing the three calibration curves for MTX, DAMPA and 7-OH-MTX used during the faecal sample analysis. Linear curve with log axis was used, all the R^2 for each of the standards for both the urine and the faecal samples were superior to 0.97.

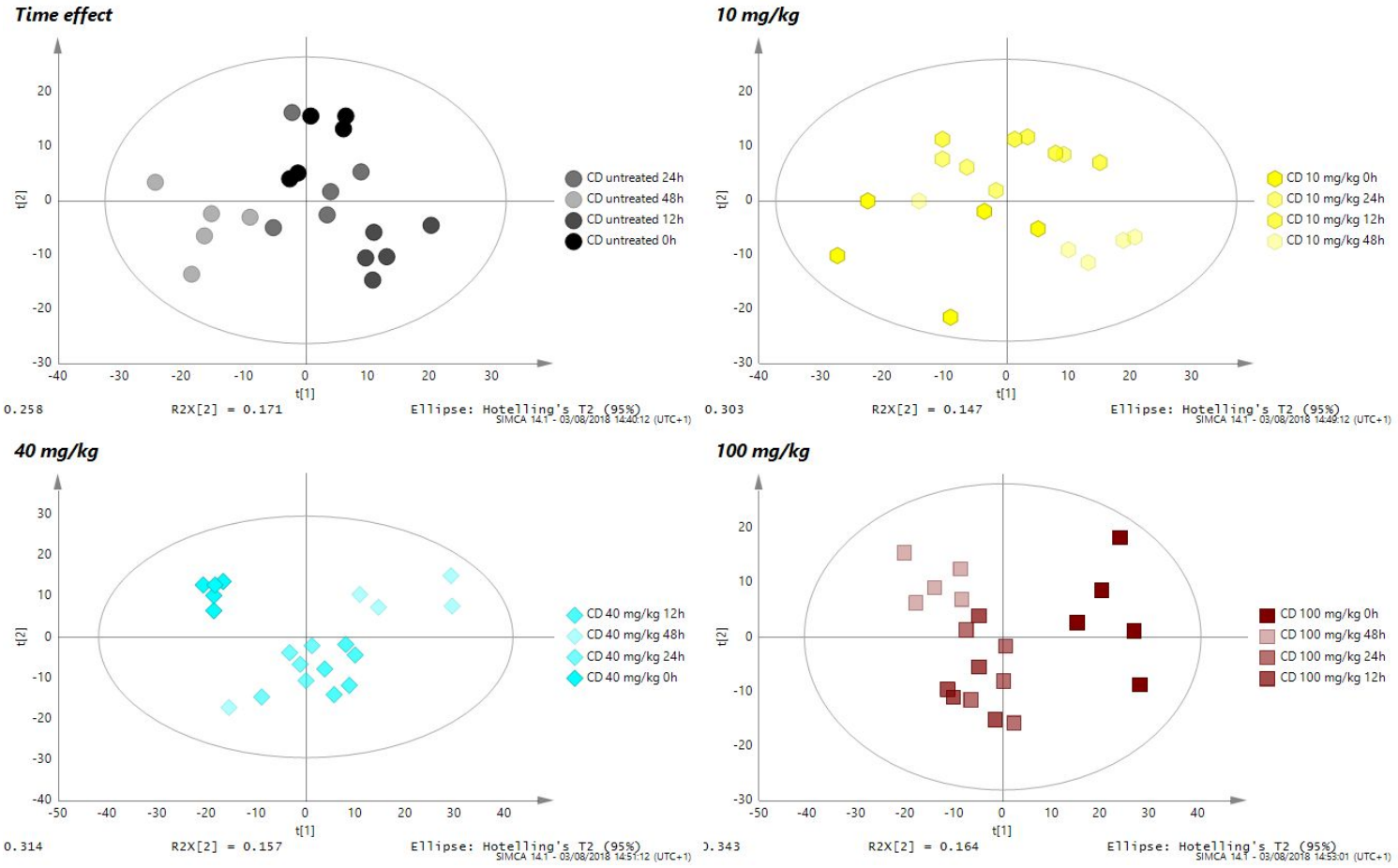
Time effect
CD untreated 24h
CD untreated 48h
CD untreated 12h
CD untreated 0h
t[1]
t[2]
0.258
R2X[2] = 0.171
Ellipse: Hotelling's T2 (95%)
SIMCA 14.1 - 03/08/2018 14:40:12 (UTC+1)
10 mg/kg
CD 10 mg/kg 0h
CD 10 mg/kg 24h
CD 10 mg/kg 12h
CD 10 mg/kg 48h
0.303
R2X[2] = 0.147
Ellipse: Hotelling's T2 (95%)
SIMCA 14.1 - 03/08/2018 14:49:12 (UTC+1)
40 mg/kg
CD 40 mg/kg 12h
CD 40 mg/kg 48h
CD 40 mg/kg 24h
CD 40 mg/kg 0h
0.314
R2X[2] = 0.157
Ellipse: Hotelling's T2 (95%)
SIMCA 14.1 - 03/08/2018 14:51:12 (UTC+1)
100 mg/kg
CD 100 mg/kg 0h
CD 100 mg/kg 48h
CD 100 mg/kg 24h
CD 100 mg/kg 12h
0.343
R2X[2] = 0.164
Ellipse: Hotelling's T2 (95%)
SIMCA 14.1 - 03/08/2018 14:53:01 (UTC+1)

Figure S2. PCA score plots of the urine metabolic profile, plotted according the dose of MTX administrated. Control are indicated by black/grey circles and 10, 40 or 100 mg/kg are indicated by yellow hexagons, blue diamonds, dark red boxes respectively. A colour intensity gradation for each colour indicates the different time-periods as indicated in the key.

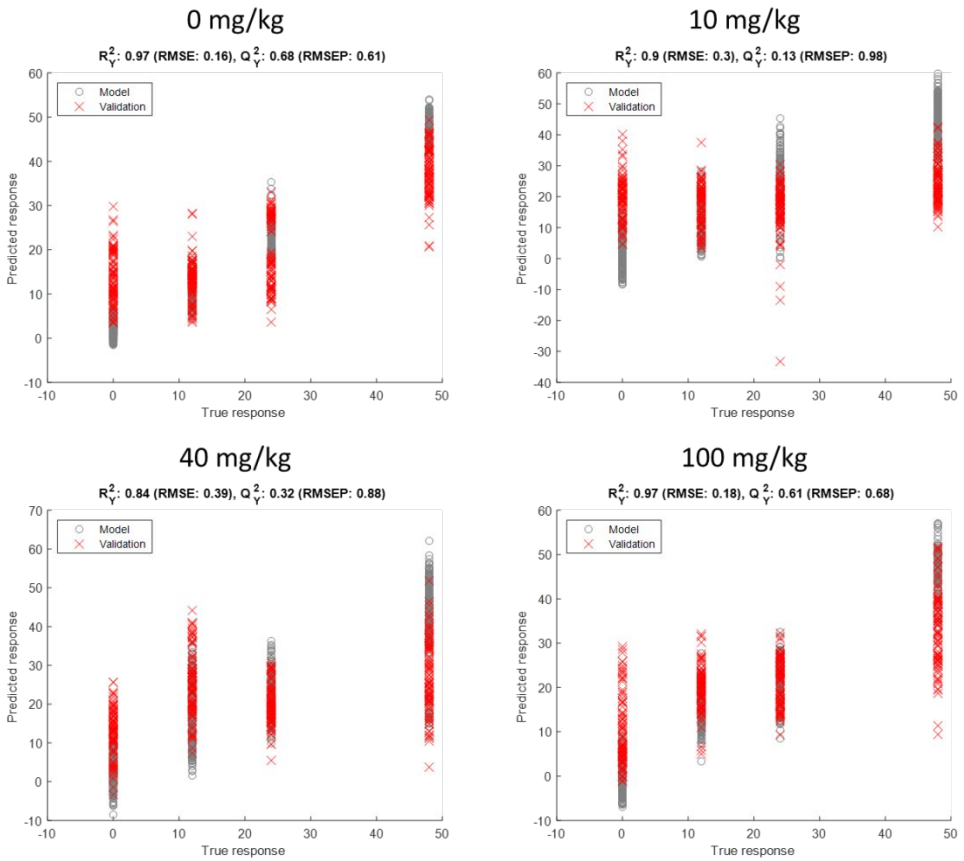

Figure S3. CA-PLS-R (covariate-adjusted projection to latent structures regression) analysis performed on the urine samples over the time points 0, 12, 24 and 48h for each of the MTX dose- groups. R2Y indicates the goodness of fit for the model of the data, the

Q2Y how well the model predicts the data. RMSE and RMSEP represent the root mean squared error and the root mean squared error of the prediction respectively. The model set is represented by grey circles and the predicted ones by red crosses.

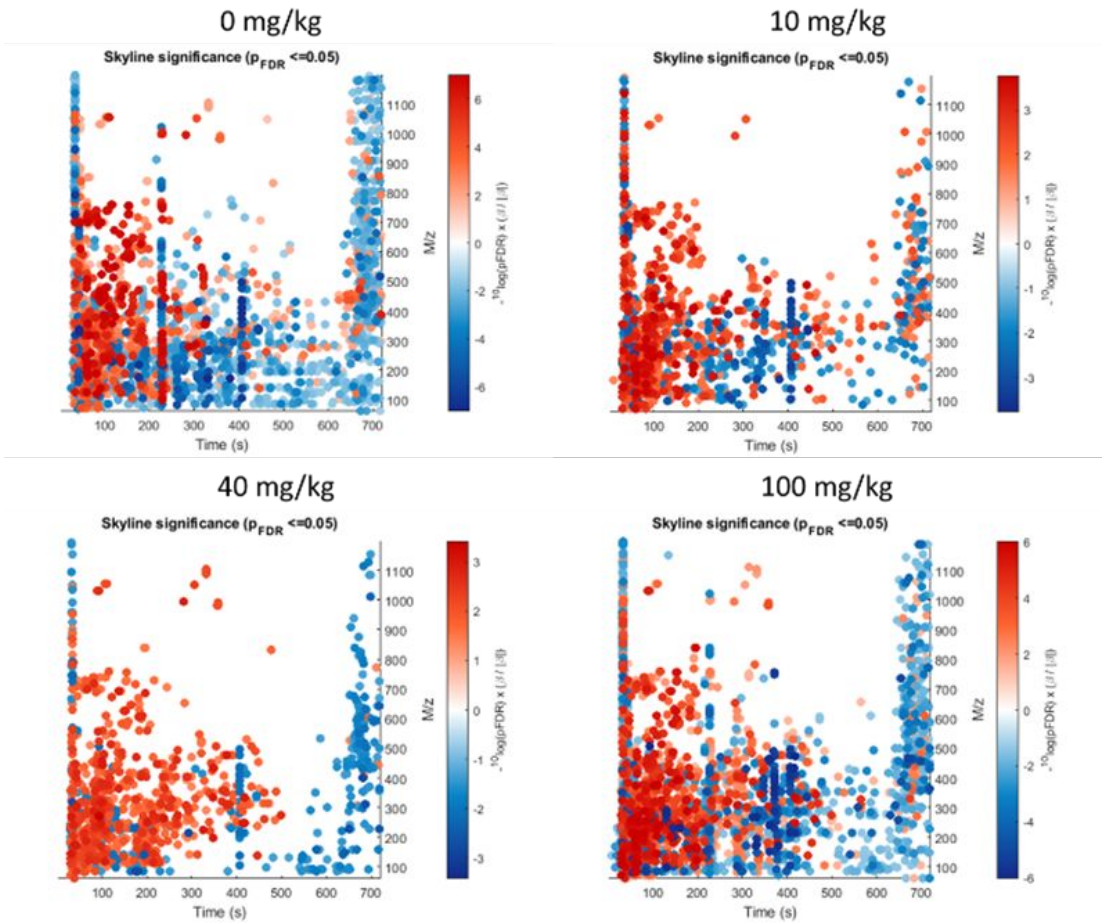
0 mg/kg
Skyline significance (p_{FDR} <=0.05)
10 mg/kg
Skyline significance (p_{FDR} <=0.05)
40 mg/kg
Skyline significance (p_{FDR} <=0.05)
100 mg/kg
Skyline significance (p_{FDR} <=0.05)
Time (s)
M/z
$-^{10}\log(pFDR) \times (\beta / |\beta|)$

Figure S4. Skyline significance of the CA-PLS-Regression models, performed on the urine samples for each of the MTX dose groups (0, 10, 40 and 100 mg/kg), representing the number of features having a q-value less than 0.05 and which significantly increased (red colour) or decreased (blue colour) over time (0, 12, 24 and 48h).

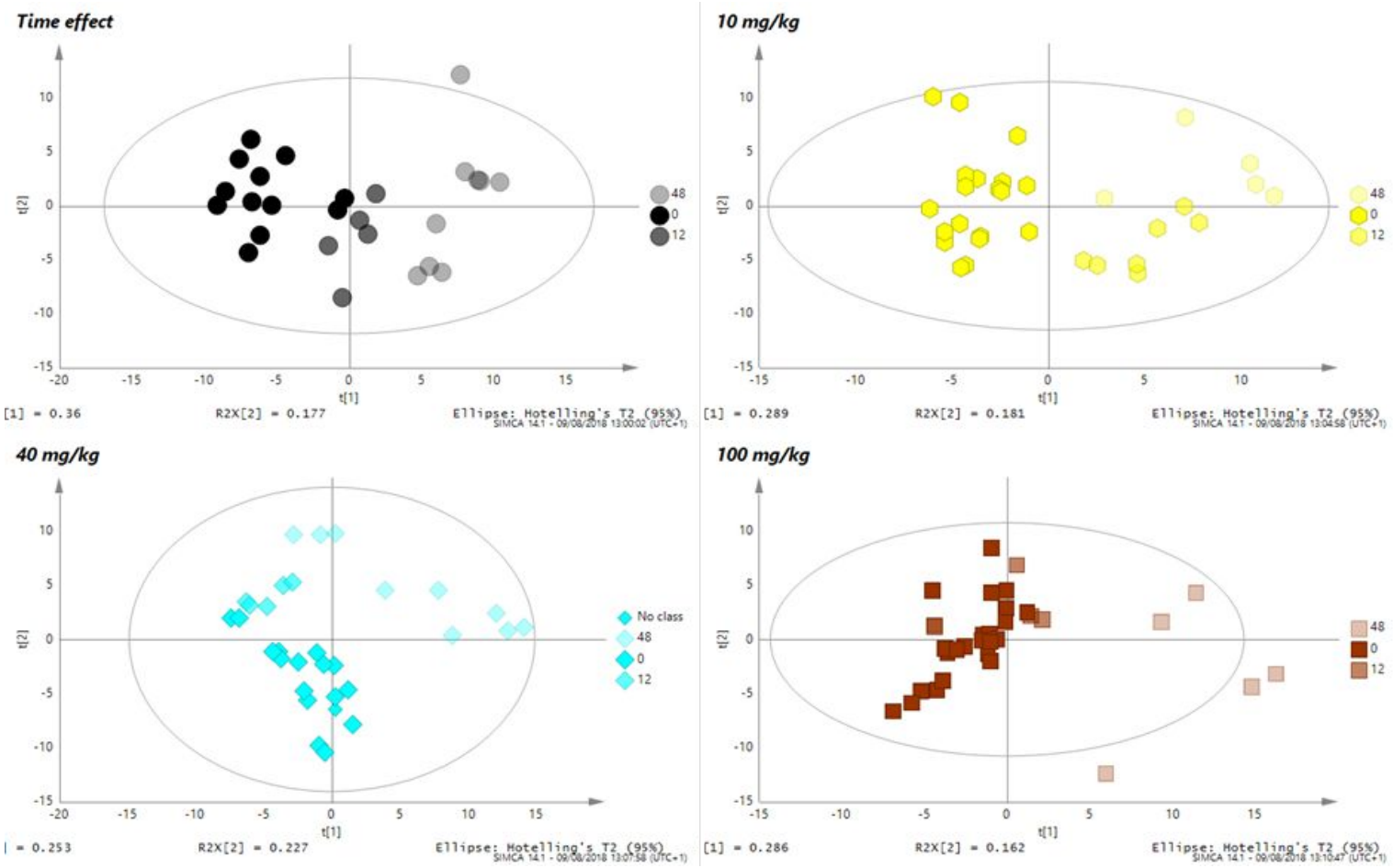

Figure S5. PCA score plots of the faecal metabolic profiles plotted according the dose of MTX administered. The undosed controls are indicated by black/grey circles and the animal groups treated with 10, 40 or 100 mg/kg are indicated by yellow hexagons, blue

diamonds, dark red boxes respectively. A colour intensity gradation for each colour indicates the different time-periods as indicated in the key.

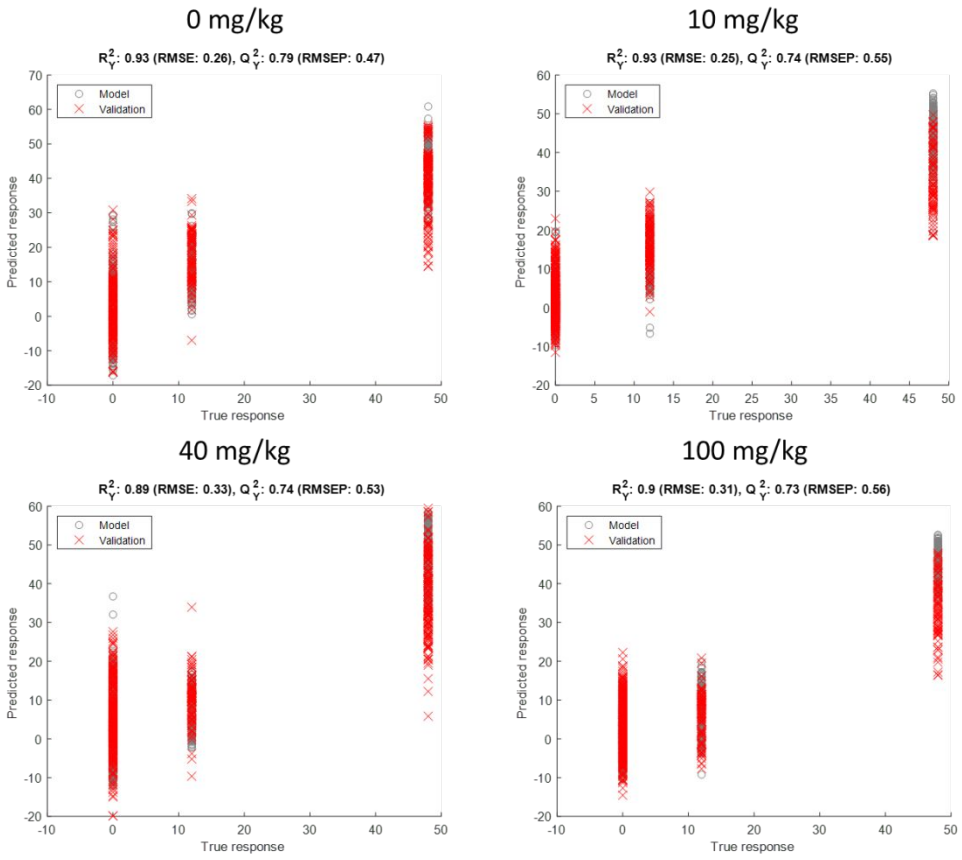

Figure S6. CA-PLS-R (covariate-adjusted projection to latent structures regression) analysis performed over time (0, 12 and 48h) on the faecal samples of each of the MTX dose groups. R2Y indicate how well the model fits the data, the Q2Y how well the model predicts the data. RMSE and RMSEP represent the root mean squared error and the root mean squared error of the prediction respectively. The model set is represented by grey circles and the predictions by red crosses.

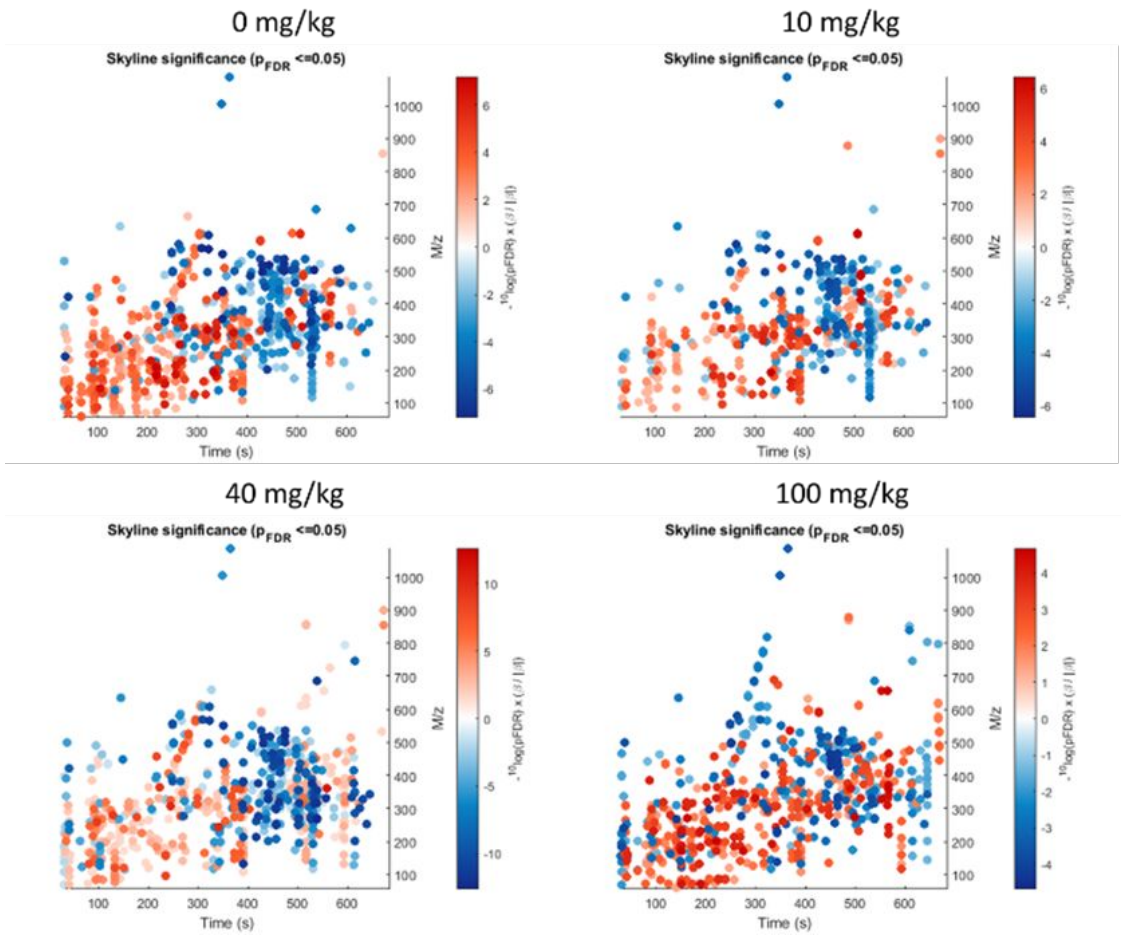
0 mg/kg
Skyline significance (p_{FDR} <=0.05)
10 mg/kg
Skyline significance (p_{FDR} <=0.05)
40 mg/kg
Skyline significance (p_{FDR} <=0.05)
100 mg/kg
Skyline significance (p_{FDR} <=0.05)
M/z
Time (s)
$-^{10}\log(pFDR) \times (sl\ |\ il)$

Figure S7. Skyline significance of the CA-PLS-Regression models, performed on faecal samples for each of the MTX dose groups (0, 10, 40 and 100 mg/kg), representing the number of features having a q-value lower than 0.05 and which significantly increased (red colour) or decreased (blue colour) over time (0, 12 and 48h).

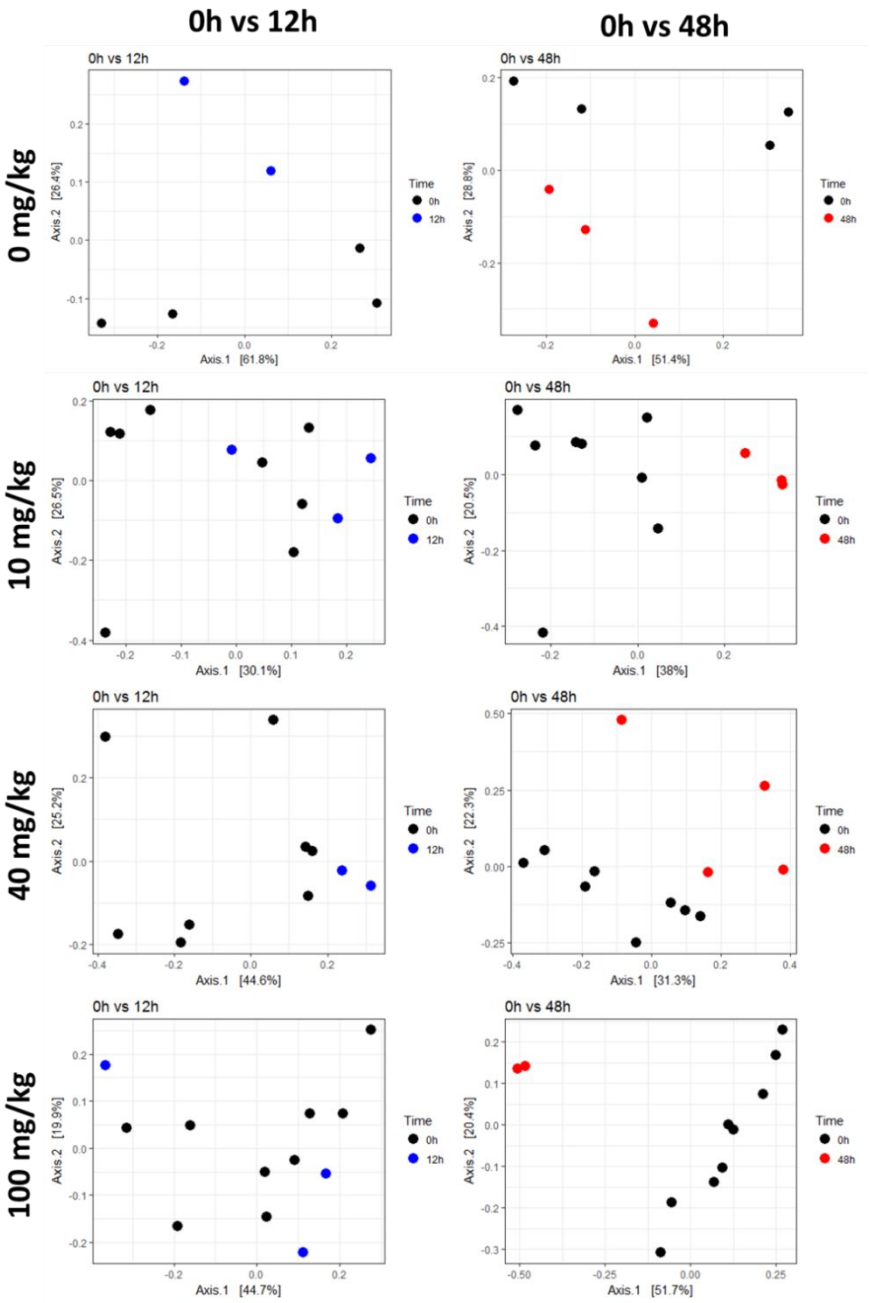

Figure S8. β diversity represented using PCoA of Bray-Curtis distance, displaying the similarity between the samples. For the untreated animals, N = 4 at 0 h, N = 2 at 12 h and N = 3 at 48 h; for the animals which received a dose of 10 mg/kg, N = 8 at 0 h, N = 3 at 12 h and N = 3 at 48 h; for the animals which received a dose of 40 mg/kg, N = 8 at 0 h, N = 2 at 12 h and N = 4 at 48 h; for the animals which received a dose of 100 mg/kg, N = 9 at 0 h, N = 3 at 12 h and N = 2 at 48 h.

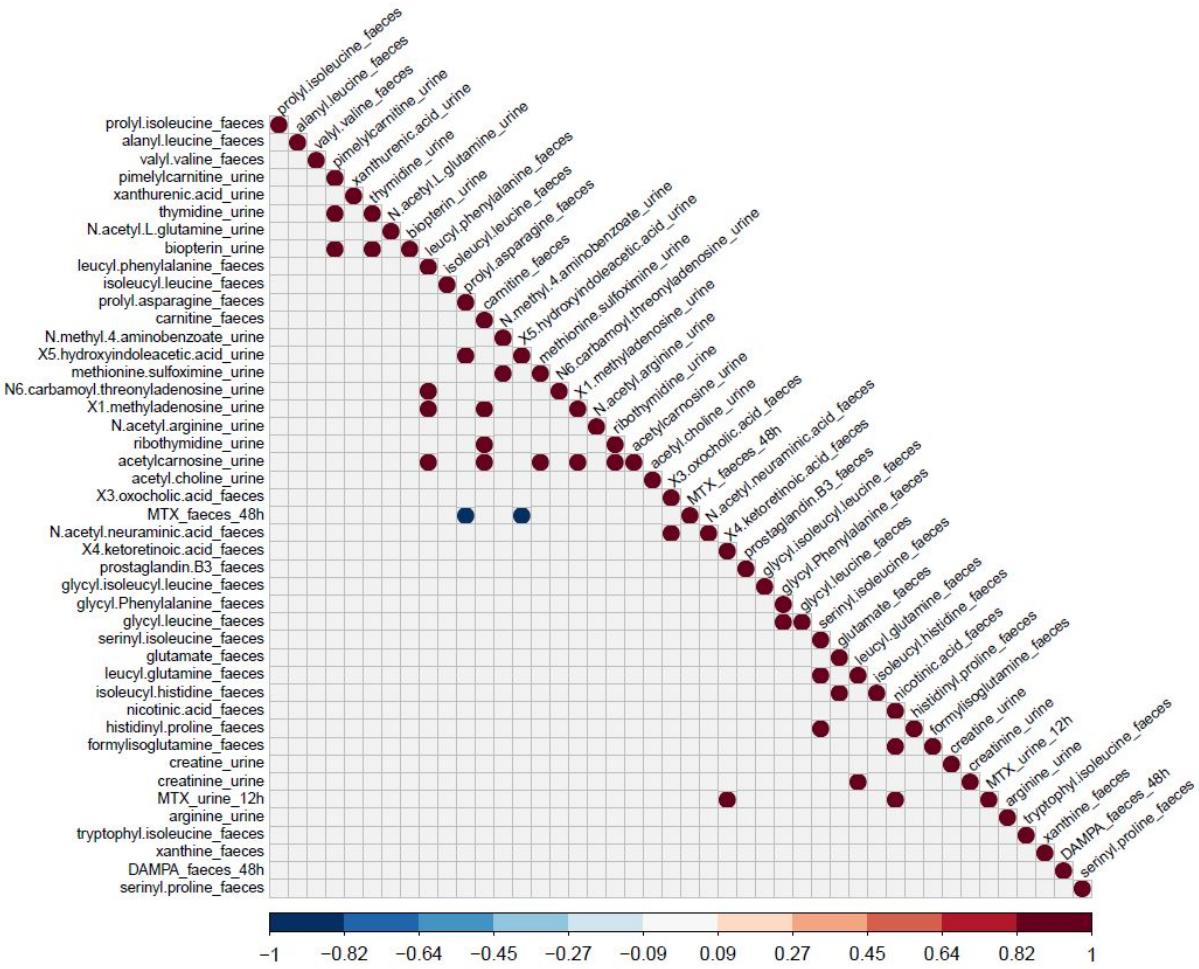

Figure S9. Correlogram integrating the putatively annotated endogenous compounds excreted in urine and feces samples of the animals pre-MTX treatment with MTX concentration measured in urine at 6-12 h and in feces at 36-48 h, as well as the quantity of DAMPA excreted in feces samples at 36-48 h. Represents the significant correlations only ($p < 0.05$) showing negative correlations for 5-hydroxyindole acetic acid in urine samples of the animals pre-MTX treatment with MTX excretion in feces at 48h. N = 3 animals.

Table S1. Number of reads per samples after 16S rRNA gene sequencing pre-processing and prior importation into phyloseq for data investigation.

SAMPLE ID	READS COUNT
10_C_01	13189
10_C_02	13680
11_C_01	13129
11_C_02	6182
11_C_04	15638
11_C_08	12801
12_C_01	16269
13_D_04	15603
14_D_01	17043
14_D_02	17909
14_D_08	14648
15_D_02	16451
16_D_02	15168
16_D_08	11959
17_E_01	12937
17_E_02	19481
17_E_08	10813
18_E_01	15943
18_E_08	14209
19_E_01	15241
19_E_08	16935
1_A_08	15753
20_E_01	19037
20_E_02	14600
21_F_01	16923
21_F_02	18220
22_F_01	18429
22_F_08	17227
23_F_01	18852
23_F_02	18097
23_F_08	18209
24_F_01	17406
24_F_02	4406
25_G_01	15517
25_G_02	16624

25_G_04	17014	33_I_08	18292
25_G_08	6745	34_I_01	18632
26_G_01	18234	34_I_02	14898
26_G_02	18370	35_I_01	11371
27_G_01	11336	35_I_02	12434
27_G_04	13117	35_I_04	17798
28_G_01	17482	35_I_08	16482
28_G_02	15805	36_I_01	13159
29_H_01	12397	36_I_02	14176
29_H_04	14809	36_I_08	17722
29_H_08	11919	37_I_01	10524
2_A_02	18056	37_I_02	13151
2_A_08	15282	38_I_01	14775
30_U_01	16765	38_I_02	13278
30_U_02	21982	39_J_02	10433
31_H_01	15485	3_A_02	11592
31_H_08	12619	40_J_01	17328
32_H_01	13381	40_J_02	15795
32_H_02	12724	41_J_01	12942
33_I_02	14062	41_J_02	14064
33_I_04	13899	42_J_01	17428

42_J_02	18626
42_J_08	19071
43_J_01	12622
43_J_04	11234
44_J_01	24371
44_J_02	14249
4_A_01	17464
4_A_02	13804
5_B_01	12107
5_B_02	15130
5_B_04	14601
5_B_08	15550
6_B_01	18907
7_B_02	15476
8_B_01	17688
8_B_02	12730
8_B_08	21445
9_C_01	11312
9_C_04	10702
9_C_08	15240
BLANK-P1 181	53
BLANK-P2 257	71
BLANK-P3-V1 248	80
BLANK-P3-V2 324	23
POSITIVE-CONTROL 30404	23749

Table S2. Putative annotations for metabolites detected in the urine samples. The database reference number of each putative annotation is indicated, as well as if the matching was done with as spectra coming from real measurements or in silico calculation, and finally if validation of the annotation by MS/MS or DDA is reported as well as matches with the NPC database. Confidence of annotation are indicated in ID level. Frag = fragment; theo = theoretical; obs = observed.

putative ID	Adduct	formula	ID level	m/z	Calculated mass	frag 1 theo	frag 2 theo	frag 3 theo	frag 4 theo	frag 1 obs	frag 2 obs	frag 3 obs	frag 4 obs	RT (mins)	Database	In silico or exp	MS/MS or DDA	NPC DB
arginine	[M+H]+	C6H14N4O2	2a	175.0828	175.1187	158.093	130.098	116.071	112.087	158.057	130.053	/	112.056	0.61	Metlin 13	ESI Q-TOF	Yes	Yes
creatinine	[M+Na]+	C4H7N3O	2a	136.0504	136.0481	114.069	86.0741	/	/	114.068	86.0722	/	/	0.65	Metlin 8	ESI Q-TOF	/	Yes
creatine	[M+H]+	C4H9N3O2	2a	132.0786	132.0768	114.066	90.0555	87.0557	/	114.068	90.0568	87.0566	/	0.67	Metlin 7	ESI Q-TOF	Yes	Yes
acetyl choline	[M+H]+	C7H15NO2	2a	146.12	146.118	87.0448	64.0165	60.0818	/	87.0459	/	60.0826	/	0.69	Metlin 57	ESI Q-TOF	/	Yes
ribothymidine	[M+Na]+	C10H14N2O6	2b	281.0753	281.0744	259.093	133.048	127.05	110.023	259.094	133.052	127.052	/	0.95	Metlin 85160	ESI Q-TOF	Yes	/
N-acetyl-L-glutamine	[M+Na]+	C7H12N2O4	2a	211.0711	211.0689	189.087	172.06	130.05	84.0444	189.089	172.061	130.052	84.045	1.07	Metlin 58344	ESI Q-TOF	Yes	Yes
acetylcarnosine	[M+Na]+	C11H16N4O4	2b	291.1070	291.1064	269.124	251.114	209.058	156.077	269.126	251.115	209.058	156.077	1.08	HMDB0012881	In silico	Yes	/
N-acetyl-arginine	[M+H]+	C8H16N4O3	2b	217.1320	217.1298	200.103	175.117	158.083	/	200.105	175.12	158.084	/	1.08	Metlin 58241	ESI Q-TOF	Yes	/
1-methyladenosine	[M+H]+	C11H15N5O4	2b	282.1216	282.1200	264.109	206.067	150.077	/	/	206.019	150.08	/	1.40	Metlin 6888	ESI Q-TOF	Yes	/
biopterin	[M+Na]+	C9H11N5O3	2b	260.078	260.075	238.094	220.083	194.068	178.073	238.096	220.085	194.069	178.073	1.58	Metlin 247	ESI Q-TOF	Yes	/
thymidine	[M+H]+	C10H14N2O5	2a	265.0830	265.0795	243.114	127.05	117.055	99.0439	243.056	127.054	117.057	/	2.25	Metlin 3375	ESI Q-TOF	/	Yes
xanthurenic acid	[M+H]+	C10H7NO4	2b	206.048	206.044	188.034	178.046	160.038	132.044	188.074	178.053	160.042	132.082	2.97	Metlin 5841	ESI Q-TOF	Yes	/
N6-carbamoyl-L-threonyladenosine	[M+Na]+	C15H20N6O8	2b	435.1255	435.1235	413.142	281.099	136.062	102.016	413.143	281.101	136.064	102.055	3.14	Metlin 95993	In silico	Yes	/
pimelylcarnitine	[M+H]+	C14H25NO6	2b	304.178	304.176	286.165	268.154	258.17	240.159	286.174	268.11	258.18	239.997	3.15	HMDB0013328	In silico	Yes	/
methionine sulfoximine	[M+H]+	C5H12N2O3S	2b	181.0630	181.0640	163.038	135.059	118.032	102.055	163.052	135.058	118.066	102.016	3.20	Metlin 65992	ESI Q-TOF	Yes	/
5-Hydroxyindoleacetic acid	[M+H]+	C10H9NO3	2b	192.0676	192.0637	146.06	110.009	/	/	146.063	110.011	/	/	3.54	Metlin 2975	ESI Q-TOF	Yes	/
N-Methyl-4-aminobenzoate	[M+H]+	C8H9NO2	2b	152.0720	152.0710	134.0600	108.0810	106.0650	104.0490	134.0616	108.0823	106.0653	104.0100	3.66	Metlin 65994	In silico	Yes	/

Table S3. Putative annotations for metabolites detected in the faecal samples. The database reference number of each putative annotation is indicated, as well as if the matching was done with as spectra coming from real measurements or in silico calculation, and finally if validation of the annotation by MS/MS or DDA is reported as well as matches with the NPC database. Confidence of annotation are indicated in ID level. **Frag = fragment; theo = theoretical; obs = observed**

putative ID	adduct	formula	ID level	*m/z*	Calculated mass	frag 1 theo	frag 2 theo	frag 3 theo	frag 4 theo	frag 1 obs	frag 2 obs	frag 3 obs	frag 4 obs	RT (mins)	Database	In silico or exp	MS/MS or DDA	NPC DB
carnitine	$[M+H]^+$	$C_7H_{15}NO_3$	2a	162.113	162.113	103.04	102.092	85.0291	60.0821	103.04	102.055	85.048	60.081	0.62	Metlin 52	ESI Qtof	Yes	Yes
prolyl-asparagine	$[M+H]^+$	$C_9H_{15}N_3O_4$	2b	230.113	230.114	213.087	195.076	133.061	98.06	213.089	195.078	133.061	98.0602	0.65	Metlin 85910	In sillico	Yes	/
N-acetyl-neuraminic acid	$[M+H]^+$	$C_{11}H_{19}NO_9$	2a	310.12	310.12	292.106	274.093	232.083	108.044	292.102	274.091	232.083	108.049	0.66	Metlin 24101	ESI Qtof	Yes	Yes
formylisoglutamine	$[M+H]^+$	$C_6H_{10}N_2O_4$	2b	175.073	175.071	157.061	147.076	139.05	130.05	157.062	147.077	139.057	130.051	0.66	Metlin 63495	In sillico	Yes	/
histidinyl-proline	$[M+H]^+$	$C_{11}H_{16}N_4O_3$	2b	253.128	253.13	235.119	138.066	116.071	110.071	235.118	138.056	116.071	110.075	0.67	Metlin 85799	In sillico	Yes	/
serinyl-proline	$[M+H]^+$	$C_8H_{14}N_2O_4$	2b	203.099	203.103	185.092	171.076	157.097	116.071	185.088	171.076	157.092	116.072	1.05	Metlin 85940	In sillico	Yes	/
nicotinic acid	$[M+H]^+$	$C_6H_5NO_2$	2b	124.04	124.039	106.029	80.0501	78.0344	/	106.048	80.05	78.034	/	1.11	Metlin 240	ESI Qtof	Yes	/
xanthine	$[M+H]^+$	$C_5H_4N_4O_2$	2b	153.041	153.04	136.013	110.035	82.0402	/	136.015	110.035	82.041	/	1.45	Metlin 82	ESI Qtof	Yes	/

isoleucyl-histidine	$[M+H]^+$	$C_{12}H_{20}N_4O_3$	2b	269.159	269.161	251.15	223.155	156.077	114.091	251.146	223.157	156.076	/	1.47	Metlin 85815	In sillico	Yes	/
leucyl-glutamine	$[M+H]^+$	$C_{11}H_{21}N_3O_4$	2b	260.16	260.16	243.134	147.077	130.05	86.0964	243.133	147.076	130.05	86.0963	1.61	Metlin 85831	In sillico	Yes	/
glutamate	$[M+H]^+$	$C_5H_9NO_4$	2b	148.06	148.061	130.05	105.954	102.055	84.0446	130.04	105.947	102.053	84.0445	1.77	Metlin 19	ESI Qtof	Yes	/
valyl-valine	$[M+H]^+$	$C_{10}H_{20}N_2O_3$	2a	217.155	217.155	199.144	171.149	118.086	100.076	199.144	171.149	118.086	/	2.09	Metlin 86032	In sillico	Yes	Yes
alanyl-leucine	$[M+H]^+$	$C_9H_{18}N_2O_3$	2a	203.139	203.142	185.131	157.136	132.104	86.0979	185.128	157.134	132.102	86.0966	2.3	Metlin 442668	ESI Qtof	Yes	Yes
serinyl-isoleucine	$[M+H]^+$	$C_9H_{18}N_2O_4$	2b	219.134	219.134	201.123	173.128	132.102	60.0444	201.124	173.129	132.102	60.0447	2.35	Metlin 85935	In sillico	Yes	/
glycyl-leucine	$[M+H]^+$	$C_8H_{16}N_2O_3$	2b	189.125	189.123	171.113	143.118	132.102	86.0964	171.113	143.119	132.103	86.0973	2.5	Metlin 5727	In sillico	Yes	/
prolyl-isoleucine	$[M+H]^+$	$C_{11}H_{20}N_2O_3$	2b	229.151	229.152	183.147	70.0645	/	/	183.143	/	/	/	2.78	Metlin 23879	ESI Qtof	Yes	/
glycyl-Phenylalanine	$[M+H]^+$	$C_{11}H_{14}N_2O_3$	2b	223.103	223.108	177.102	166.086	120.081	103.054	177.1	106.081	120.077	103.053	2.9	Metlin 85758	In sillico	Yes	/
isoleucyl-leucine	$[M+H]^+$	$C_{12}H_{24}N_2O_3$	2b	245.186	245.186	227.175	199.18	132.102	114.091	227.176	199.18	132.102	114.097	3.45	Metlin 85817	In sillico	Yes	/
glycyl-isoleucyl-leucine	$[M+H]^+$	$C_{14}H_{27}N_3O_4$	2b	302.205	302.204	171.109	143.114	132.099	/	171.107	143.113	132.095	/	3.76	Metlin 23190	ESI Qtof	Yes	/
leucyl-phenylalanine	$[M+H]^+$	$C_{15}H_{22}N_2O_3$	2b	279.170	279.17	233.165	132.102	120.081	103.054	233.155	132.102	120.081	103.055	4.12	Metlin 85839	In sillico	Yes	/

tryptophyl-isoleucine	$[M+H]^+$	$C_{17}H_{23}N_3O_3$	2b	318.182	318.181	301.155	272.176	159.092	132.102	301.156	272.182	159.093	132.103	5.02	Metlin 85979	In sillico	Yes	/
prostaglandin B3	$[M+H]^+$	$C_{20}H_{28}O_4$	2b	333.202	333.205	315.195	197.184	171.112	157.099	315.198	297.176	171.112	157.1	6.81	Metlin 36190	ESI Qtof	Yes	/
4-ketoretinoic acid	$[M+H]^+$	$C_{20}H_{26}O_3$	2b	315.197	315.196	297.185	269.191	251.18	149.097	297.183	269.19	251.179	/	8.36	HMDB0006285	In sillico	Yes	/
3-oxocholic acid	$[M+H]^+$	$C_{24}H_{38}O_5$	2b	407.281	407.279	389.269	371.258	347.258	343.263	389.269	371.26	/	343.264	8.62	Metlin 57908	In sillico	Yes	/

Table S4. α diversity values representing the microbial richness of each of the faecal samples of the animals according the MTX dose groups over time. The diversity is assessed using the Inverse Simpson index. Low numbers of samples due to diarrhoea and other limitations due to study design do not allow statistical tests between the α diversity of the different groups to be performed.

		Time point		
	MTX dose	0h	12h	48h
	0	14.94	14.47	15.39
	10	15.63	19.36	10.56
	40	12.30	9.66	11.18
	100	11.08	7.53	4.94
Number of samples	0	4	2	2
	10	5	3	2
	40	7	1	3
	100	6	2	2