Gut Microbiota Dynamics And Uremic Toxins: One Size Does Not Fit All

SUPPLEMENTARY MATERIALS AND METHODS

Study population

Eighteen CKD patients, treated with post-dilution hemodiafiltration (n=17) or hemodialysis (n=1; patient 6) were recruited at the dialysis unit of the Ghent University Hospital, Belgium. Exclusion criteria were active infection, active malignancy, pregnancy, unstable condition, vascular access problems, and age below 18 years. The local Ethics Committee approved the study (Ref 2012/063, B670201214999) and all patients gave written informed consent. During the 4-month study period (from September 2013 to January 2014) the data were collected.

Sample collection

Of the 18 HD patients, blood samples were collected from the vascular access prior to the mid-week dialysis session at the first study visit [week (W) 0; all in September 2013) and at the subsequent visits after 1, 2, 3, 4, 8, 12 and 16 (W1-W16) weeks. Blood samples were allowed to clot at room temperature for 20-30 min before centrifugation (10 min, 1250*g* at 4°C), serum was aliquoted and stored at -80°C until batch analysis. In parallel, fecal samples were collected. Seventeen HD patients provided one or more fecal samples that were used for microbiome analysis (Supplementary Table S1). As one of these patients (patient 13) was only able to provide a single fecal sample (W3), the total number of patients with consecutive samples was 16. The overall compliance to the fecal sampling protocol was high, with 70% of patients providing at least 7 (out of 8) fecal samples at the time of blood sampling. Upon collection, fecal samples were kept cool and they were aliquoted and stored at -80 °C within 6 hours, until batch analysis.

Quantification of uremic blood metabolites

Urea, creatinine and phosphorus were measured by standard laboratory methods. The protein-bound solutes p-cresyl glucuronide (pCG), hippuric acid (HA), indole acetic acid (IAA), indoxyl sulfate, p-cresyl sulfate and 3-carboxy-4-methyl-5-propyl-2-furanpropanoic acid (CMPF) were determined by high-performance liquid chromatography (HPLC) as previously described[1, 2, 3].

As the conversion of *p*-cresol to either *p*CG or *p*CS is covered through enzymatic conjugations by the host[4], the measurements of *p*CG and *p*CS were summed to *p*-cresyl conjugates (*p*C) when assessing their correlation to the microbial composition of the host.

Illumina-based microbial profiling

DNA was extracted from frozen fecal samples using the RNeasy PowerMicrobiome® Kit (QIAGEN group) according to the manufacturer's instructions, with minor adaptations. Bacterial and archaeal 16S rRNA genes were amplified using the 515F/806R primer set, targeting the V4 hypervariable region[5]. Sequencing was performed using the Illumina MiSeq platform with sequencing kit MiSeq v2, producing 250 bp paired-end reads. Sequence analysis was performed according to Geirnaert *et al.*[6]. Using UCHIME, FLASH-merged sequences were subjected to quality filtering with the FASTX-toolkit and chimera removal[7, 8]. Samples were rarefied and taxonomical classifications of sequences were performed with RDP classifier[9] to generate phylum to genus level composition matrices. Bootstrap values from the RDP classifier were used to identify sequences with high-confidence genus assignments

(bootstrap value >0.8), while sequences classified with lower confidence were binned on family level (labeled unclassified family). A species-level matrix was generated by de novo OTU clustering with the UPARSE pipeline[10] using 97% identity threshold as described earlier[11]. Genus abundances were rarefied to 9,000 reads per sample.

Microbiota data visualization and analyses

Triplot

A triplot with Bray Curtis as dissimilarity measure on genus abundances was carried out using vegan. The top 10 taxa that most strongly co-varied with overall community composition and all significant non-redundant numeric metadata identified with envfit are added as arrows, of which the lengths represent the relative strength of the co-variation with overall community composition.

PCoA plot

The PCoA plots were visualized using R packages ggplot2[12] and phyloseq[13].

Redundancy analysis (RDA)

RDA first expresses variables as a linear combination of underlying factors (multiple regression step). Then it carries out a PCoA on these linear combinations of underlying factors. Here, we selected two toxins (*p*C and IxS) as variables to test to what extent they could be explained by the taxon counts as underlying factors. Unclassified taxa and taxa with a prevalence below 20 were summed into two separate dummy taxa and then removed. Taxa were centred and scaled (mean zero and standard deviation one).

Comparison of inter- versus intra-patient beta-diversity

Intra- versus inter-patient beta-diversity was assessed by *i)* computing all pair-wise community dissimilarity values with Bray Curtis within each patient and for all patients and testing whether they were significantly different with the unpaired two-sided Wilcoxon rank sum test and *ii)* fitting a Dirichlet-Multinomial distribution with R package dirmult[14], thereby quantifying the overdispersion θ.

Linear discriminant analysis (LDA) effect size (LEfSe)

LEfSe was run on the Galaxy server with default settings (http://www.bioinformatics.nl/galaxy/), except for disabling normalization, since the data were already rarefied to the same total sample count.

Statistical analysis

Patient characteristics

The descriptive statistical analysis on the ESKD patient characteristics was performed with SPSS Statistics 24 (2016, Armonk, NY, USA). Data were checked for normality using the Shapiro-Wilk test. Due to missing data specific tests were used to analyze the data, for comparisons of the different time points of normally distributed paired data a linear mixed model for repeated measures followed by Bonferroni's multiple testing correction was used. For this, the different time points and each patient were set respectively as repeated and subject variables. Next, the different time points were introduced as a fixed factor and the numeric data as dependent variables. For comparisons between different time points of not-normally distributed paired data, Skillings-Mack test (package Skillings.Mack) was performed[15]. This statistical test was performed with R (2016, Version 3.3.1.). P-values < 0.05, or in case of multiple testing correction q-values < 0.05, were considered significant.

Data processing

Anatomical therapeutic chemical (ATC) classification system codes were merged up to the second level, including information on the anatomical main group and therapeutic subgroup. Metadata for which more than half of the values were missing and constant metadata were discarded, leaving 92 binary, 27 categorical and 46 numeric metadata items.

For envfit, metadata were further processed to remove all missing values. For this, all metadata items with more than 4 missing values in all patients at all time-points combined, were discarded. Remaining missing values were replaced with patient-specific means for numeric metadata items or the most frequent patient-specific categorical/binary value for categorical/binary metadata items. The missing-value free metadata set contains 56 binary, 13 categorical and 31 numeric metadata items.

Selection of metadata items co-varying with community composition

A principal coordinates analysis (PCoA) with Bray Curtis as dissimilarity measure on rarefied genus abundances using vegan was carried out[16]. Vegan's function envfit was then run on the PCoA results and on missing-value-free metadata with 1000 permutations. All metadata items with Benjamini-Hochberg-corrected envfit p-values below 0.05 were deemed to significantly correlate with community composition. To ensure that results are robust to small noise, genus abundances were re-rarefied independently 10 times and only those metadata items that were significant across all iterations were retained. To reduce collinearity, all-pairwise Pearson correlations between envfit-selected metadata were computed.

For categorical data, a Pearson correlation analysis of the curated metadata was performed (threshold: 0.7). The correlating parameters formed clusters, represented as edges between metadata items in a graph and identified groups as connected components in this graph using R package igraph[17]. Per cluster the variable with the highest effect size (i.e. highest envfit score) was retained as a non-redundant covariate, with the exception of toxins, for which each group member was retained (see Supplementary Table S2). This way, we identified 36 metadata items that co-vary significantly with overall community composition (Supplementary Figure S3).

SUPPLEMENTARY RESULTS

Patient characteristics

Characteristics of the 17 ESKD patients are summarized in Supplementary Table S3. Their mean age was 73 ±10 years. The main causes of CKD were diabetic nephropathy (53%), renal vascular disease (24%) and glomerulonephritis (18%). No significant differences were found between different time points (Supplementary Table S3). Eight ESKD patients were hospitalized during the study period for the following reasons: surgical intervention for arterio-venous fistula stenosis (patient (p)1, p16), dialysis catheter replacement (p6), surgery for pseudo-aneurysm in the groin (p10), fistula infection (p15, p16), bronchitis exacerbation (p17), finger amputation (p11), dyspnea and orthopnea (p11), renal cyst bleeding (p6), cholecystitis (p18) and unstable angina (p10). All but p1 and p16 received antibiotic treatment during their admission. Patient 7 was treated with antibiotics at W1 for bronchitis and patient 6 had recurrent liver cyst infection, resulting in a need for preventive treatment with antibiotics (temocillin, a β-lactamase-resistant penicillin).

Average residual renal function, calculated as the arithmetic mean of creatinine and urea clearance, was 4.62 ml/min and was significantly anti-correlated to serum values of IxS, IAA, HA and creatinine but not to the other metabolites measured (data not shown).

Characterization of the gut microbiome in ESKD

Co-variates of intestinal microbiota composition in ESKD

Envfit combined with clustering was used to identify non-redundant parameters that significantly correlated with the overall microbiota composition, through correlation with PCoA eigenvectors. Length of scaled arrows reflects correlation with overall community composition. Supplementary Figure S3 depicts co-variates of gut microbiota composition as selected by Envfit with the individual patient as main co-variate followed by medical history, medication use, blood values, Bristol stool score (BSS), anthropometric values and lifestyle parameters.

Comparison of BSS between patients on antibiotics versus not on antibiotics was assessed according to the unpaired two-sided Wilcoxon rank sum test. The test was repeatedly (100 times) applied to randomly selected sub-sets of non-antibiotic samples to compare an equal number of non-antibiotic and antibiotic samples. No significant difference in BSS between patients on antibiotics versus not on antibiotics was observed.

The correlation between intestinally generated uremic toxins and transit time was assessed, by looking at their correlation with BSS. The strongest correlation was found between IAA and BSS, followed by IxS and HA (Spearman correlation: 0.40, 0.32 and 0.30, respectively). As part of the samples were taken from participants on laxatives, the analyses were also repeated excluding samples of patients on laxatives, to avoid data distortion due to laxative usage (Supplementary Table S4) and the correlation between IAA and BSS became stronger (Spearman correlation: 0.50).

Variation in microbiota profiles in ESKD

The composition of fecal microbiota of patients in the present cohort was compared with that of subjects of the Flemish Gut Flora Project (FGFP)[11] which have a similar genetic and environmental background. The microbiome of the ESKD patients was compared to both the whole population-based cohort of 1106 subjects as well as to a subset of age-matched non-diseased controls (n=32). A significantly higher variability of the gut microbiome was observed in the 17 ESKD patients as compared to average subject-to-subject differences, both for the population-based cohort ($P<0.0001$) and the subset ($P<0.0001$; Supplementary Figure S1).

Projecting the ESKD patients' samples on the PCoA plot of the FGFP confirmed that these patients do not cluster in a specific area but rather are dispersed over the entire space of the control population. This indicates that the composition of gut microbiota is diverse among ESKD patients and that there is no overall shared compositional signature in this cohort (Supplementary Figure S2).

Uremic toxin concentrations and associations with gut microbiota composition

Associations between IxS and *p*C and specific microbial taxa

Using ordistep (with 500 permutations and 50 steps), we tried to assess which linear combination of taxa could significantly explain the observed metabolite concentrations. However, no significant explanatory linear combination of taxa could be identified.

Next, the fecal microbiota of patients with highest *p*C and lowest IxS serum concentrations (pC_{high}_IxS_{low} dataset) in this cohort was compared to the fecal microbiota of patients with highest IxS and lowest *p*C serum concentrations (pC_{low}_IxS_{high} dataset) to identify potential microbial targets for concentration management. Using the quantile function in R, we selected the samples in the intersection of lowest and highest quantiles, i.e. [0.7,1] and [0,0.3] quantiles, for both toxins. Five samples resorted in the pC_{high}_IxS_{low} dataset and 15 samples fulfilled the criteria of the pC_{low}_IxS_{high} dataset. Since the two groups had different sample numbers, also sub-sampling was tested.

The Dirichlet-Multinomial (DM) distribution was fitted to each group with the HMP R package in order to extract the taxon proportion vector, which is a parameter of the DM distribution. Assuming that taxon counts are well described by the DM, the two taxon proportion vectors were significantly different according to Generalized Wald test statistics ($p<0.0001$).

We performed a RDA using fecal bacterial taxa to try to explain the *p*C and IxS blood values, whereby the two selected metabolites were expressed as a linear combination of taxa (Supplementary Figure S4). Using this approach, the proportion of metabolite variance explained by the genera is 69%. Based on the PCoA plot, it is apparent that *Faecalibacterium* and *Bacteroides* point towards high IxS and *Lactobacillus*, *Enterococcus* and *Ruminococcus* tend to be abundant in patients with high *p*C blood values (Supplementary Figure S4).

Characterization of the stability of gut microbiota in ESKD patients over time

Correlating uremic toxins with overall within-patient gut microbiota community changes

The longitudinal sampling allowed us to assess if changes in microbiota composition correlated with changes in uremic toxin concentrations. Therefore, differences in uremic toxin concentrations were compared to the Bray Curtis dissimilarity indices for each sample pair within one patient. Changes in *p*C and IxS correlated significantly with community changes in the corresponding samples from the same patients (Supplementary Figures S5A and B). Also, IAA and creatinine variations correlated significantly with overall microbiota compositional changes (data not shown). No correlations were found here between changes in all other blood metabolites and fecal microbial community dissimilarity.

SUPPLEMENTARY REFERENCES

1 Eloot S, Van Biesen W, Roels S, Delrue W, Schepers E, Dhondt A*, et al.* Spontaneous variability of pre-dialysis concentrations of uremic toxins over time in stable hemodialysis patients. *PLoS One* 2017;**12**:e0186010.

2 Fagugli RM, De Smet R, Buoncristiani U, Lameire N, Vanholder R. Behavior of non-protein-bound and protein-bound uremic solutes during daily hemodialysis. *Am J Kidney Dis* 2002;**40**:339-47.

3 Meert N, Schepers E, Glorieux G, Van Landschoot M, Goeman JL, Waterloos MA*, et al.* Novel method for simultaneous determination of p-cresylsulphate and p-

cresylglucuronide: clinical data and pathophysiological implications. *Nephrol Dial Transplant* 2012;**27**:2388-96.
4 Gryp T, Vanholder R, Vaneechoutte M, Glorieux G. p-Cresyl Sulfate. *Toxins* 2017;**9**.
5 Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J, Fierer N*, et al.* Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME J* 2012;**6**:1621-4.
6 Geirnaert A, Wang J, Tinck M, Steyaert A, Van den Abbeele P, Eeckhaut V*, et al.* Interindividual differences in response to treatment with butyrate-producing Butyricicoccus pullicaecorum 25-3T studied in an in vitro gut model. *FEMS Microbiol Ecol* 2015;**91**.
7 Magoc T, Salzberg SL. FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* 2011;**27**:2957-63.
8 Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R. UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* 2011;**27**:2194-200.
9 Wang Q, Garrity GM, Tiedje JM, Cole JR. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl Environ Microbiol* 2007;**73**:5261-7.
10 Edgar RC. UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nat Methods* 2013;**10**:996-8.
11 Falony G, Joossens M, Vieira-Silva S, Wang J, Darzi Y, Faust K*, et al.* Population-level analysis of gut microbiome variation. *Science* 2016;**352**:560-4.
12 Wickham H. Elegant Graphics for Data Analysis. New York: Springer-Verlag, 2009.
13 McMurdie PJ, Holmes S. phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLoS One* 2013;**8**.
14 Tvedebrink T. Overdispersion in allelic counts and -correction in forensic genetics. *Theor Popul Biol* 2010;**78**:200-10.
15 Chatfield M, Mander A. The Skillings-Mack test (Friedman test when there are missing data). *Stata J* 2009;**9**:299-305.
16 Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlinn D*, et al.* vegan: Community Ecology Package. R package version 2.4-3. 2017.
17 Csardi G, Nepusz T. The igraph software package for complex network research. *Int J* 2006;**Complex Systems**:1695.

SUPPLEMENTARY TABLES

Supplementary Table S1: Fecal sampling scheme with the number of consecutive samples per patient.

Patient number	Sample at W0	Sample at W1	Sample at W2	Sample at W3	Sample at W4	Sample at W8	Sample at W12	Sample at W16	Total number of samples per patient
1	1	1	1	1	1	1	1	0	7
2	0	0	0	0	0	0	0	0	0
3	1	1	1	1	1	1	1	1	8
4	1	1	1	1	1	0	1	1	7
5	1	1	1	1	1	1	1	0	7
6	1	1	1	0	0	1	1	1	6
7	1	1	1	1	1	1	1	1	8
8	1	1	1	1	1	1	0	1	7
9	1	0	1	1	1	1	1	1	7
10	1	1	1	1	1	1	1	1	8
11	1	1	1	1	0	0	0	1	5
12	1	1	1	1	1	1	1	1	8
13	0	0	0	1	0	0	0	0	1
14	1	1	1	1	0	1	1	1	7
15	1	1	0	0	0	1	1	0	4
16	1	1	1	1	0	1	1	1	7
17	0*	1	0	1	1	1	1	1	6
18	1	1	1	1	1	1	1	1	8
Total number of samples per time-point	15	15	14	15	11	14	14	13	111

W= week

* Sample lost due to technical reasons (insufficient number of sequencing reads)

Supplementary Table S2: Effect size of different metadata parameters and clustering based on Pearson correlation analysis.

Component	Member(s) per component (Pearson correlation >0.7)	R2 envfit score (effect size)
1	Vintage months	0.071
2	Bristol stool score	0.255
3	Target weight	0.152
	Body Mass Index	0.150
	Hip circumferences	0.105
	Weight	0.143
4	Systolic blood pressure before HD	0.103
	Systolic blood pressure	0.103
5	Indoxyl sulfate	0.210
	Creatinine	0.278
6	Indole acetic acid	0.142
7	3-carboxy-4-methyl-5-propyl-2-furanpropanoic acid	0.25
8	Hippuric acid	0.082
9	Phosphorus	0.080
10	Shannon (alpha diversity)	0.311
	Sheldon (evenness)	0.145
11	Chao1 (richness)	0.207
12	*p*-cresyl conjugates	0.16
13	Patient	0.86
	Zip code	0.774
	Residence	0.848
14	Cause of renal failure	0.356
15	Laxatives	0.047
	Type of laxative	0.395
	Laxatives use past week	0.050
	Other laxatives	0.302
	A06 (drugs for constipation)	0.047
16	Antibiotics	0.109
	Antibiotic use in the past 02 months or since last visit	0.058
	Type of antibiotic	0.124
	J01 (systemic antibacterials)	0.059
17	Prebiotics	0.171
18	Pet at home	0.061
	Cat at home	0.061
19	Dog at home	0.044
	D07 (corticosteroids, dermal)	0.070
20	Type of diabetes mellitus	0.066
21	Surgery	0.520
22	Pollen	0.067
	N05 (psycholeptics)	0.037
23	Allergy	0.067
24	B01 (antithrombotics)	0.034
25	A11 (vitamins)	0.067
26	C10 (lipid modifying agents)	0.0613

27	C07 (beta blocking agents)	0.061
28	C03 (diuretics)	0.194
29	A12 (mineral supplements)	0.092
30	C01 (cardiac therapy)	0.170
31	H02 (corticosteroids for systemic use)	0.043
32	C09 (agents acting on renin-angiotensin system)	0.041
33	B05 (blood substitutes and perfusion solutions)	0.06
34	A02 (drugs for acid related disorders)	0.035
35	G04 (urologicals)	0.051

Pearson correlation analysis of the curated metadata was performed (threshold: 0.7). 35 connected components were identified. Variable with the highest effect size (i.e. highest envfit score) per component was retained as a non-redundant covariate for further analyses.

Supplementary Table 3: ESKD patient characteristics during the study period (n = 17)

Characteristics	Time point								
	W0								
General									
Mean age in years (range)	73 ± 10 (48-87)								
Gender (M/F)	15/2								
Months on dialysis (months)	53.35 ± 29.74								
Residual renal function (ml/min)	4.62 ± 4.06								
Origin of kidney disease n (%)									
ADPKD	1 (6)								
Analgesic drugs	1 (6)								
Diabetic nephropathy	9 (53)								
Glomerulonephritis	3 (18)								
Hypertensive nephropathy	1 (6)								
Renal vascular disease	4 (24)								
Unknown	1 (6)								
	W0	**W1**	**W2**	**W3**	**W4**	**W8**	**W12**	**W16**	**q-value**
Health parameters									
Body mass index (kg/m^2)	25.13 ± 2.87	25.17 ± 2.85	25.14 ± 2.87	25.12 ± 2.86	25.59 ± 2.57	25.55 ± 2.65	25.18 ± 3.17	25.15 ± 3.10	ns[a]
Blood pressure diastolic (mmHg)	73.00 (58.00-78.50)	68.50 (54.50-83.00)	58.50 (33.50-70.50)	64.00 (52.00-82.00)	62.00 (53.00-72.00)	70.50 (54.00-74.75)	62.00 (53.50-76.00)	65.00 (61.00-74.25)	ns[b]
Blood pressure systolic (mmHg)	140.29 ± 22.15	142.12 ± 22.29	142.63 ± 16.60	140.13 ± 18.85	139.67 ± 22.63	146.69 ± 25.59	139.00 ± 16.36	151.56 ± 22.57	ns[a]
Bristol stool score	3.50 (2.00-4.00)	4.00 (3.00-6.00)	4.00 (2.75-6.00)	4.00 (2.00-6.00)	4.00 (2.50-5.00)	4.00 (3.25-4.50)	4.00 (1.75-4.50)	6.00 (3.50-6.00)	ns[b]
Dialysis parameters									
Blood flow rate (ml/min)	330.0 (300.0-347.5)	350.0 (305.0-350.0)	320.0 (300.0-350.0)	350.0 (300.0-350.0)	310.0 (250.0-350.0)	315.0 (292.5-350.0)	300.0 (295.0-345.0)	300.0 (270.0-315.0)	ns[b]
Dialyate flow rate (ml/min)	550.0 (500.0-600.0)	525.0 (506.3-600.0)	500.0 (500.0-600.0)	530.0 (500.0-600.0)	526.5 (500.0-600.0)	530.0 (500.0-600.0)	527.5 (455.0-600.0)	503.5 (460.0-600.0)	ns[b]
Kt/V	1.51 ± 0.34	No data	1.61 ± 0.23	No data	No data	1.74 ± 0.29	1.60 ± 0.26	1.52 ± 0.29	ns[a]
Ultrafiltration rate (ml/min)	1973.06 ± 1061.58	1865.47 ± 941.36	1895.50 ± 981.81	1973.93 ± 1006.69	1853.40 ± 906.36	2004.00 ± 943.03	1989.29 ± 982.52	1651.12 ± 843.02	ns[a]
Blood parameters									
Creatinine (mg/dl)	7.21 ± 2.03	7.13 ± 2.07	6.83 ± 1.99	6.96 ± 1.91	6.90 ± 1.90	7.05 ± 1.96	7.07 ± 1.97	6.99 ± 1.83	ns[a]
Phosphorus (mmol/l)	1.35 ± 0.32	1.39 ± 0.35	1.33 ± 0.25	1.36 ± 0.34	1.37 ± 0.33	1.44 ± 0.23	1.38 ± 0.26	1.42 ± 0.22	ns[a]
Urea (mg/dl)	92.7 (81.3-118.7)	98.1 (84.9-114.2)	92.1 (79.3-107.7)	99.7 (82.7-110.8)	98.6 (79.0-106.1)	92.3 (87.9-117.9)	93.9 (81.6-113.4)	89.8 (83.4-113.2)	ns[b]
C-reactive protein (mg/l)	4.50 (1.70-13.05)	No data	No data	No data	5.40 (2.35-17.95)	6.90 (1.25-18.40)	6.60 (3.55-19.25)	6.50 (1.89-22.60)	ns[b]
Uremic toxins									
CMPF (mg/dl)	0.29 (0.10-0.48)	0.30 (0.10-0.49)	0.32 (0.10-0.49)	0.25 (0.09-0.46)	0.24 (0.09-0.49)	0.28 (0.09-0.51)	0.37 (0.07-0.48)	0.40 (0.11-0.49)	ns[b]
Hippuric acid (mg/dl)	2.55 (1.03-5.38)	2.16 (0.77-6.16)	2.15 (0.89-4.59)	2.38 (0.71-3.80)	2.96 (0.81-4.29)	2.62 (0.74-4.83)	2.76 (0.89-5.85)	2.45 (0.91-4.84)	ns[b]
Indole acetic acid (mg/dl)	0.21 (0.10-0.33)	0.14 (0.10-0.28)	0.15 (0.09-0.26)	0.14 (0.09-0.26)	0.14 (0.10-0.27)	0.14 (0.10-0.29)	0.16 (0.09-0.23)	0.15 (0.08-0.20)	ns[b]
Indoxyl sulfate (mg/dl)	2.12 ± 1.25	2.20 ± 1.38	1.99 ± 1.13	1.86 ± 1.05	1.93 ± 1.26	1.97 ± 1.20	2.07 ± 1.16	1.89 ± 1.08	ns[a]
p-Cresyl glucuronide (mg/dl)	0.37 (0.18-1.05)	0.38 (0.11-1.01)	0.50 (0.12-1.04)	0.45 (0.08-0.96)	0.32 (0.13-1.11)	0.57 (0.16-1.14)	0.46 (0.14-0.90)	0.34 (0.14-0.85)	ns[b]
p-Cresyl sulfate (mg/dl)	3.19 ± 1.35	2.64 ± 1.15	2.93 ± 1.07	2.91 ± 1.30	2.76 ± 1.27	3.29 ± 1.59	3.21 ± 1.58	3.08 ± 1.88	ns[a]
Medication and supplement intake [n (%)]									
Antibiotics	1 (6)	2 (12)	1 (6)	1 (6)	2 (12)	4 (24)	5 (29)	5 (29)	
Laxatives	2 (12)	3 (18)	4 (24)	3 (18)	5 (29)	3 (18)	3 (18)	4 (24)	
Probiotics	1 (6)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	

Data presented as mean (±SD), median (25th percentile-75th percentile) or number [n (%)]. HD: Hemodialysis; ADPKD: Autosomal dominant polycystic kidney disease; Kt/V: Dialysis efficiency; CMPF: 3-carboxy-4-methyl-5-propyl-2-furanpropanoic acid; W0: First visit; W1: Second visit one week later; W2: Third visit two weeks later; W3: Fourth visit three weeks later; W4: Fifth visit four weeks later; W8: Sixth visit eight weeks later; W12: Seventh visit twelve weeks later; W16: Eighth visit sixteen weeks later; [a]: Groups differences tested with linear mixed model for repeated measures; [b]: Group differences tested with Skillings-Mack test; q-values obtained using Benjamini-Hochberg.

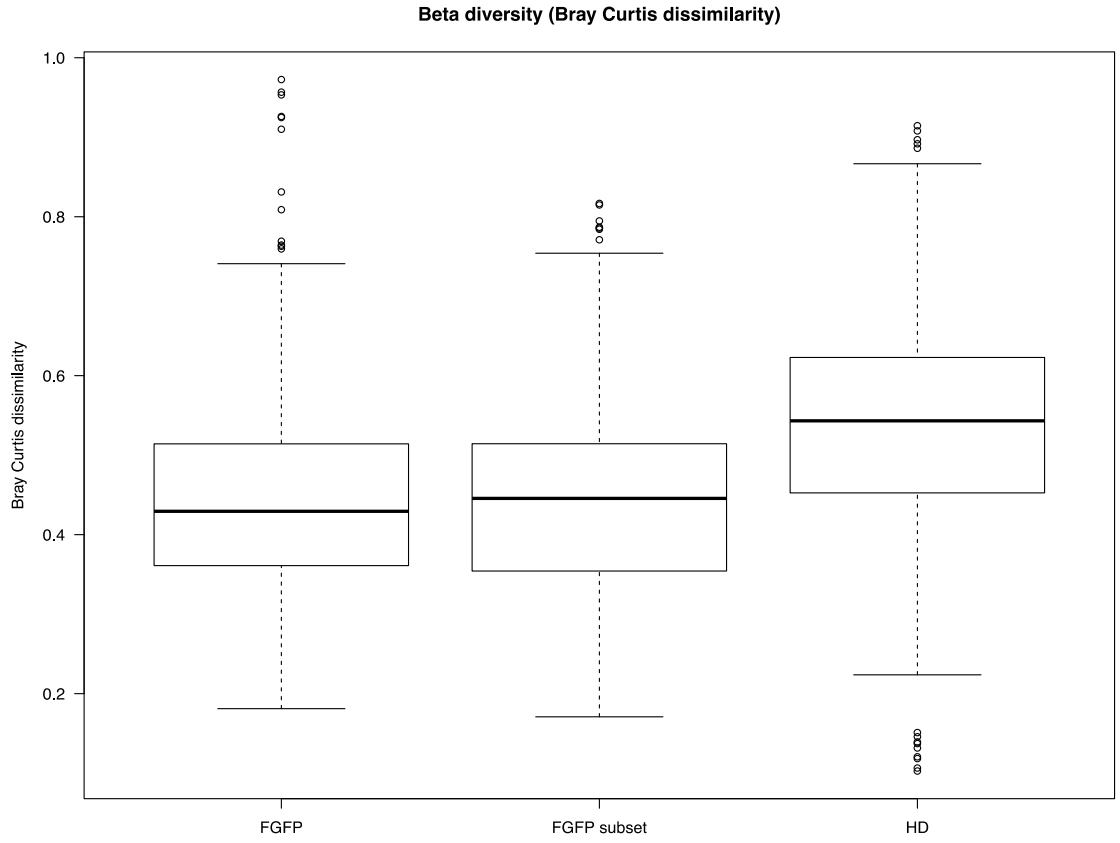
Supplementary Table 4: Correlation between intestinally generated uremic toxins and transit time

Correlation to BSS	**All samples (n=110)**		**Samples of patients not on laxatives (n=84)**	
Toxin	**Spearman correlation coefficient**	**q-value**	**Spearman correlation coefficient**	**q-value**
Indoxyl sulfate	**0.32**	**0.0055**	**0.28**	0.09
Indole acetic acid	**0.40**	**0.0001**	**0.50**	**1.60 e-06**
Hippuric acid	**0.30**	**0.0140**	**0.23**	ns
p-Cresyl conjugates	0.10	ns	-0.09	ns

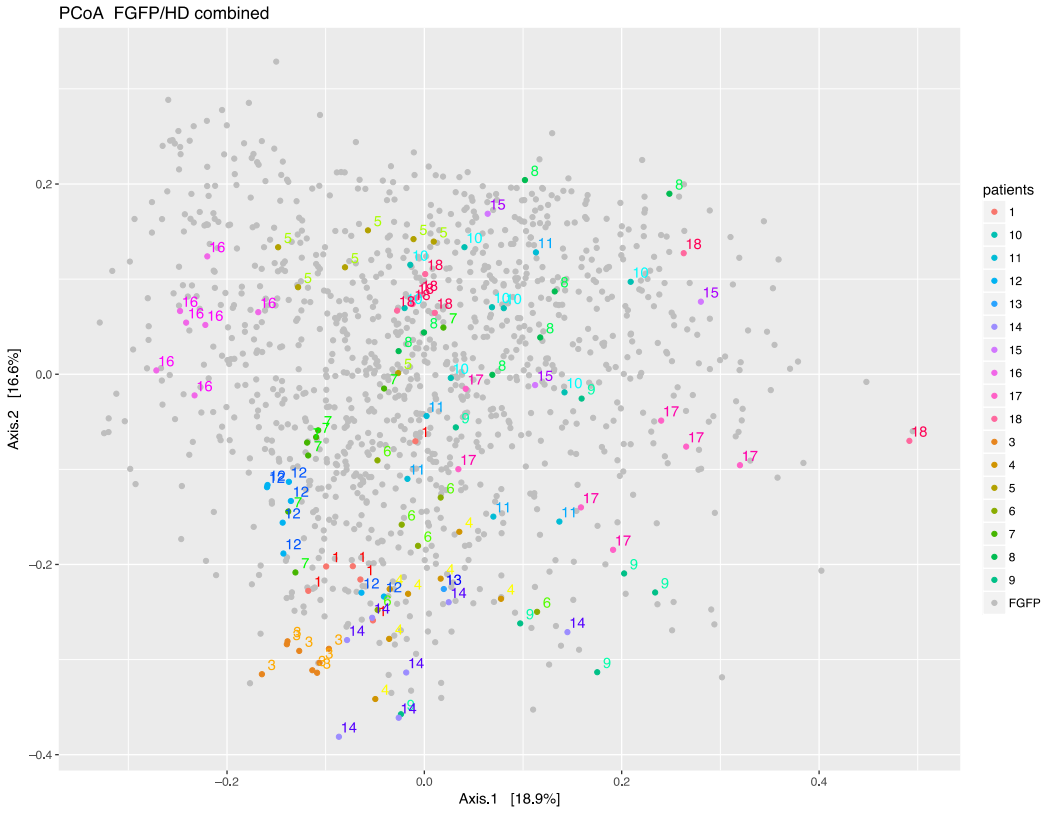
BSS= Bristol stool score which is used to assess transit time (slow transit= low BSS). Bonferroni correction was applied.

SUPPLEMENTARY FIGURES

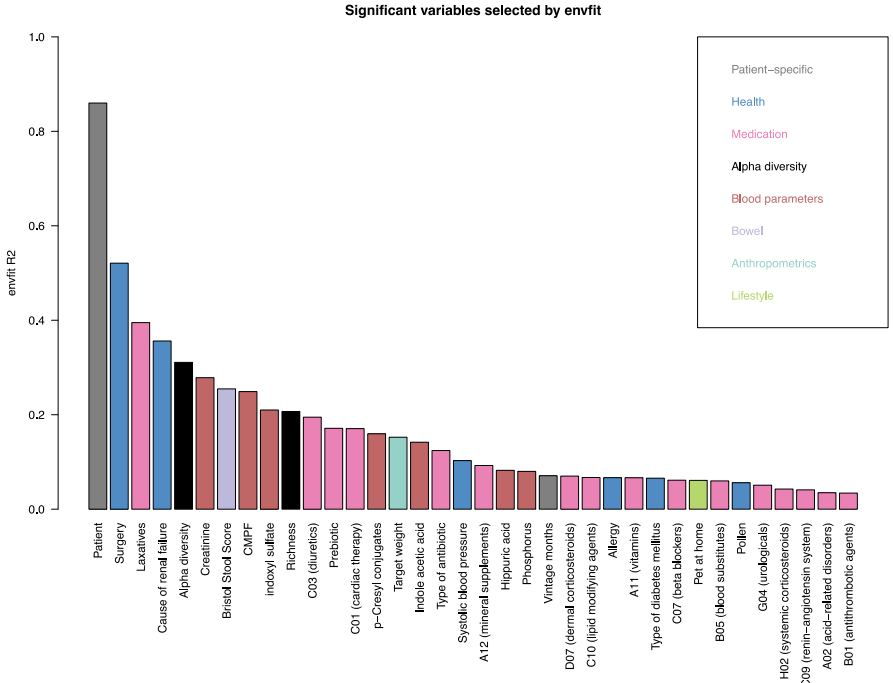
Supplementary Figure S1



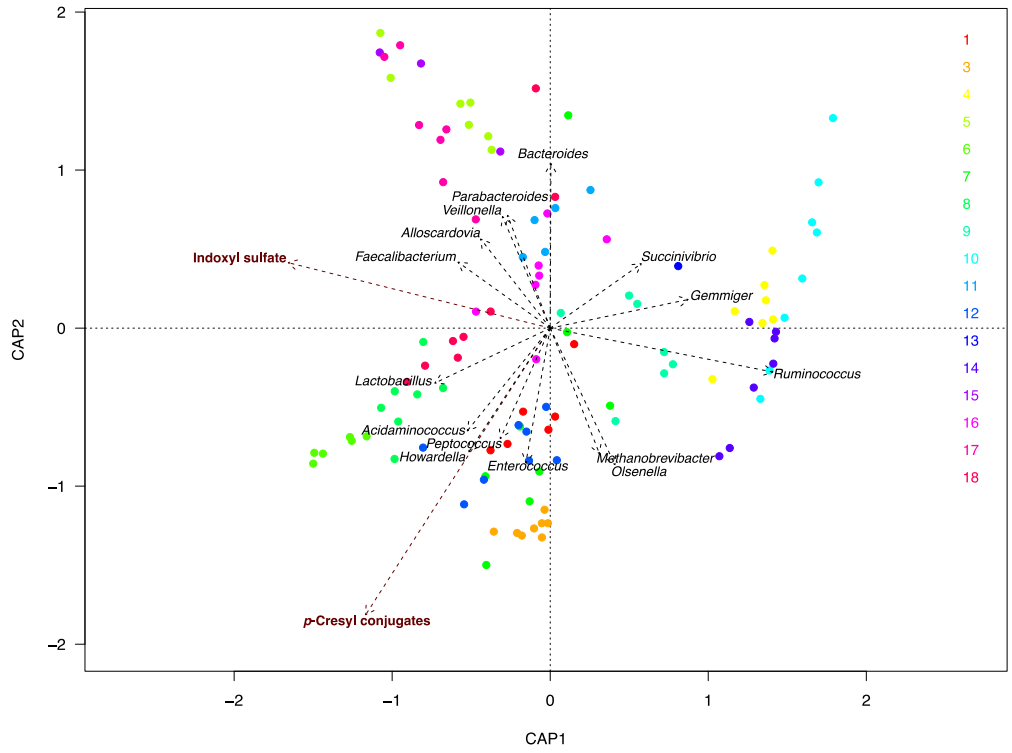
Supplementary Figure S2



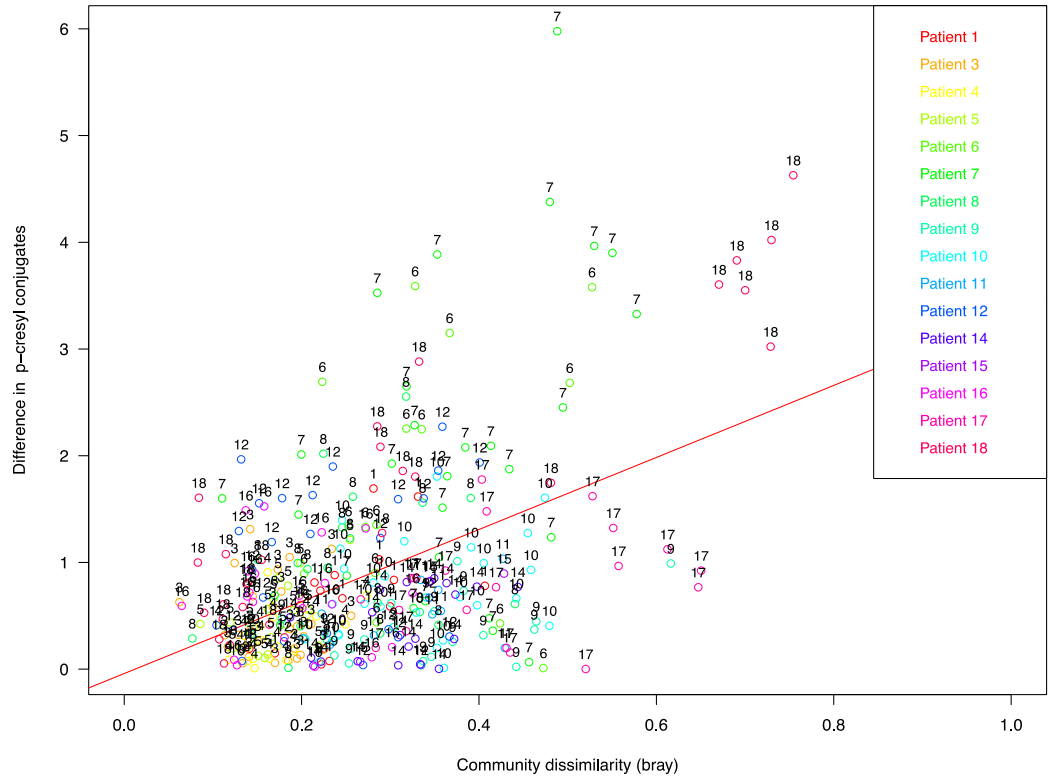
Supplementary Figure S3



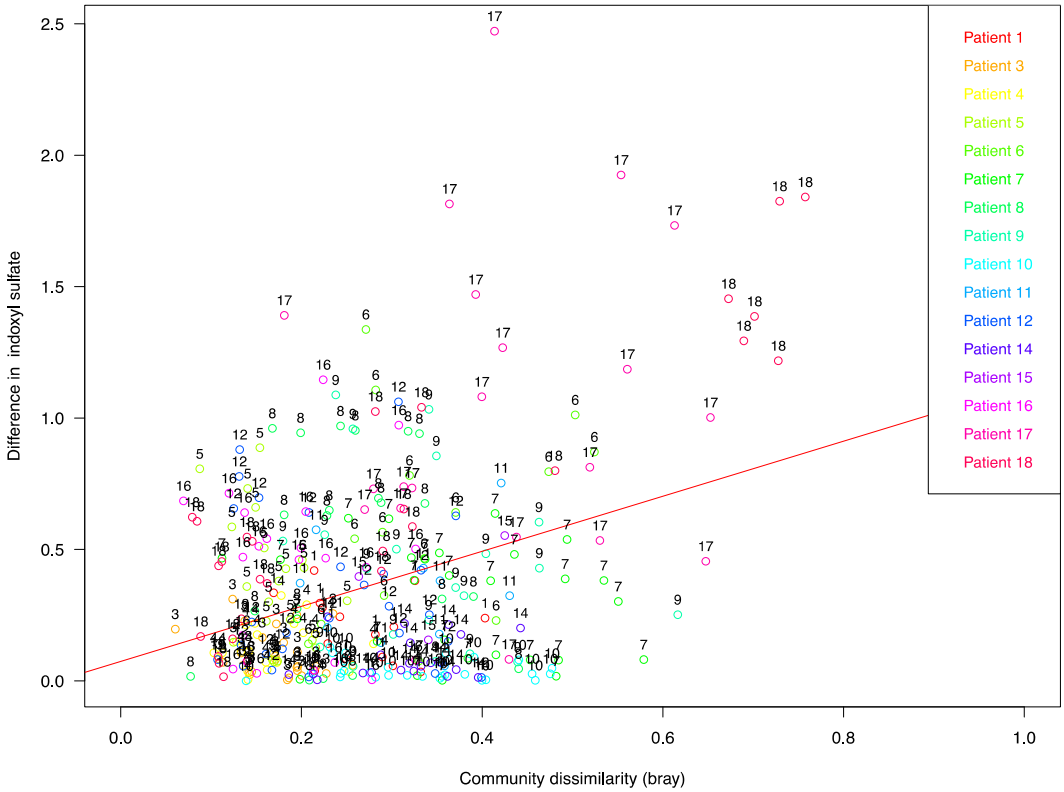
Supplementary Figure S4



Supplementary Figure S5A



Supplementary Figure S5B



LEGENDS TO SUPPLEMENTARY FIGURES

Supplementary Figure S1.
The gut microbiome of ESKD patients is more variable than that of average subjects. The box plots illustrate the span of all pair-wise dissimilarities in each of the three data sets (Flemish Gut Flora Project (FGFP), healthy age-matched FGFP and ESKD). To adjust for the different number of sample pairs, the number of sample pairs in the smallest data set (496 in the age-matched FGFP) was randomly selected from the two other data sets. The dissimilar distribution of the ESKD data has a significantly larger mean (p-value < 0.0001) than both the FGFP and age-matched FGFP data sets according to the Wilcoxon rank sum test, whereas there is no significant difference between the means of the FGFP and FGFP age-matched data sets.

Supplementary Figure S2.
Combined PCoA plot depicting samples from an average population (FGFP; in gray) and all samples collected in the present study (other colors - genera not present in both data sets removed). Dissimilarity computed with Bray Curtis.

Supplementary Figure S3.
Main co-variates of the fecal microbiota composition of ESKD patients. Final list of metadata selected by envfit, colored according to the Flemish Gut Flora Project (FGFP)[11] metadata categories (pink: medication; red: blood parameters; purple: bowel; dark blue: health; light blue: anthropometrics; green: life style). CMPF: 3-carboxy-4-methyl-5-propyl-2-furanpropanoic acid. Alpha diversity was calculated using the Shannon index as the sum of weighted taxon proportions. Richness equals the number of different taxa in a sample and was estimated using the Chao1 index. Larger

values of the Shannon and Chao1 index represent higher alpha diversity and richness, respectively. Medication was classified according to the ATC-codes.

Supplementary Figure S4.
Proportion of *p*-cresyl conjugates and IxS blood values variance explained by fecal bacterial genera. A redundancy analysis (RDA) using the fecal bacterial taxa to explain the *p*-cresyl and IxS blood values was done. On the PCoA, *p*C and IxS metabolite variance is expressed as a combination of fecal bacterial genera.

Supplementary Figure S5.
Correlation between changes in blood metabolites and changes in fecal microbial composition. Intra-patient community dissimilarity (Bray Curtis) change of fecal microbiota versus uremic toxin concentration change for the concurrent blood sample pairs depicted for A. *p*-cresyl conjugates (R^2= 0.236; $p<1.10^{-5}$) and B. indoxyl sulfate (R^2= 0.122; $p<1.10^{-5}$).