Metabonomics of human fecal extracts characterize ulcerative colitis, Crohn's disease and healthy individuals

Jacob Tveiten Bjerrum• Yulan Wang•
Fuhua Hao• Mehmet Coskun• Christian Ludwig•
Ulrich Gunther• Ole Haagen Nielsen

Metabolomics (2015) 11:122–133

DOI 10.1007/s11306-014-0677-3

Metabonomics

- Approach pioneered by Jeremy Nicolson at Imperial College London
- Defined as “the quantitative measurement of the dynamic multiparametric metabolic response of living systems to pathophysiological stimuli or genetic modification”
- Metabonomics extends metabolic profiling to include study of perturbations caused by extragenomic sources such as environment, microflora, and disease processes

J K Nicolson et al (1999, 2002, 2006)

Introduction

- Ulcerative colitis (UC) and Crohn’s disease (CD) - two distinct but similar disease – are part of Inflammatory Bowel Disease (IBD) syndrome
- Caused by unbalanced immunological response to the intestinal microbes and its components in genetically predisposed individuals

Objective of the study

- Current diagnosis approach for IBD involves clinical history, endoscopy, radiology, histology, microbiology, and hematology
- Invasive, time consuming and still leaves around 10% patients as unclassified
- To develop non invasive metabonomics approach to identify metabolic markers from fecal extracts to differentiate between IBD patients and controls as well as between the IBD subtypes

Materials and Methods

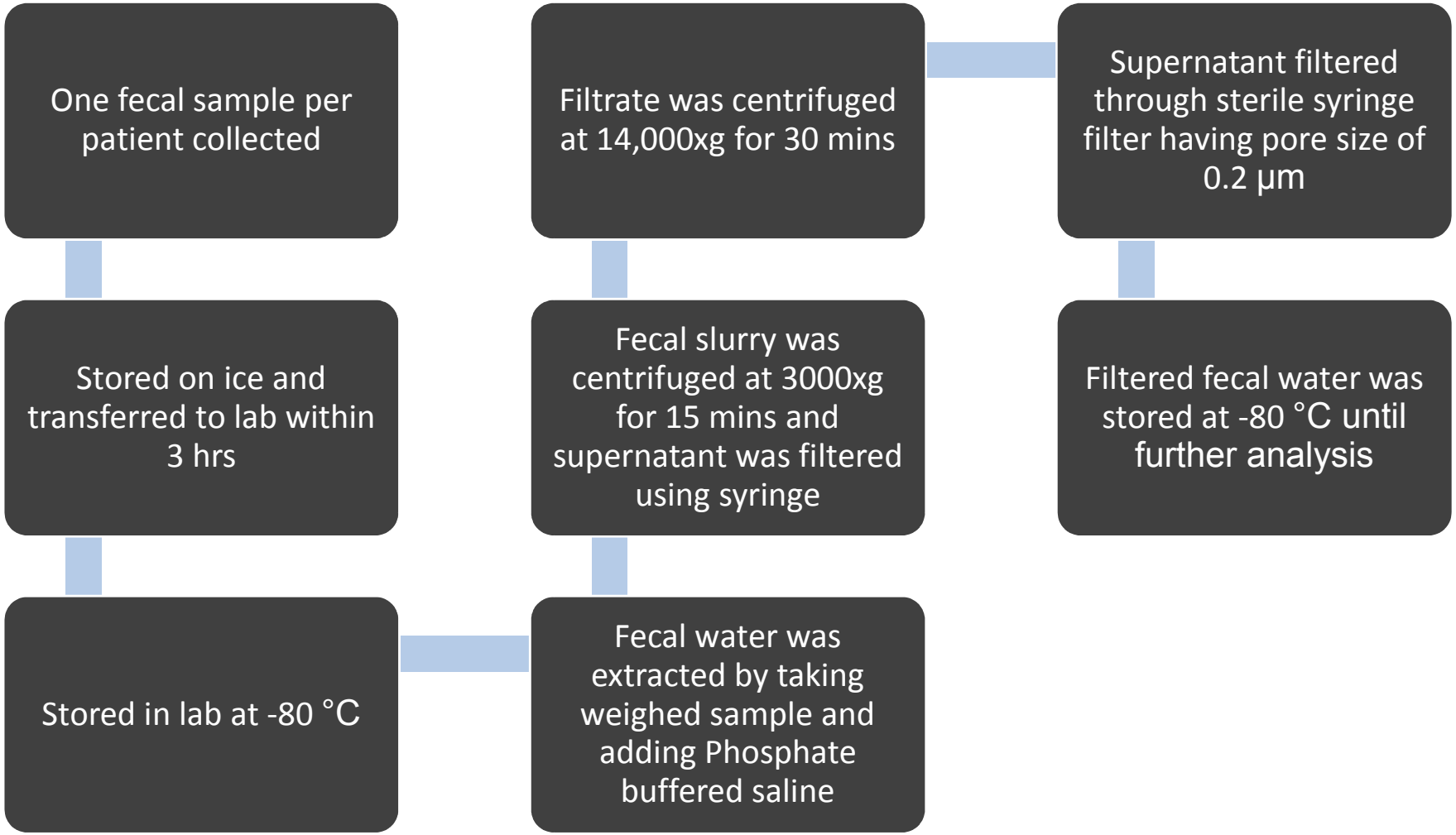
Table 1 Clinical details

Characteristics	Inactive CD n = 31	Active CD n = 13	Inactive UC n = 29	Active UC n = 19	Controls n = 21
Gender (male/female)	13/18	6/7	14/15	11/8	13/8
Age, years (mean, range)	44 (18–80)	36 (22–57)	48 (27–72)	48 (25–80)	40 (18–73)
Age at diagnosis (≤25/>25 years)	10/21	8/5	5/24	3/16	–
Years with disease (≤10/>10 years)	16/15	9/4	12/17	14/5	–
HB-score (mean, range)	1 (0–4)	10 (5–18)	–	–	–
Mayo-score (mean, range)	–	–	0.3 (0–1)	7 (3–11)	–
Extension (P/PS/LC/PC/Ileit)	–	2/2/3/1/9	–	1/2/4/11/0	–
Surgery (IR/IR+HC/HC/C)	6/2/2/3	3/2/1/1	0/0/1/2	0/0/0/0	–
Smoking/non-smoking	9/22	5/8	7/22	2/17	2/19
EIM (present/not present)	6/25	1/12	1/28	1/18	–
Steroids, n					
Responder/non-responder/unknown	24/0/7	9/3/1	17/2/10	14/3/2	–
Independent/dependent/unknown	21/3/7	4/8/1	17/2/10	10/7/2	–
Daily medication, n					
Sulfasalazine (1.5–2.0 g)	1	0	2	1	–
Systemic 5-aminosalicylic acid (1.6–3.2 g)	2	0	24	17	–
Topical 5-aminosalicylic acid (1 g)	0	0	5	11	–
Systemic glucocorticoids (75 mg)	1[a]	4	2[a]	10	–
Topical glucocorticoids (100 mg)	0	1	1	3	–
Azathioprine (100–150 mg)	9	3	2	0	–
6-Mercaptopurine (50–75 mg)	1	0	1	0	–
Infliximab (5 mg/kg)	3	4	1	0	–
None	17	3	2	0	21

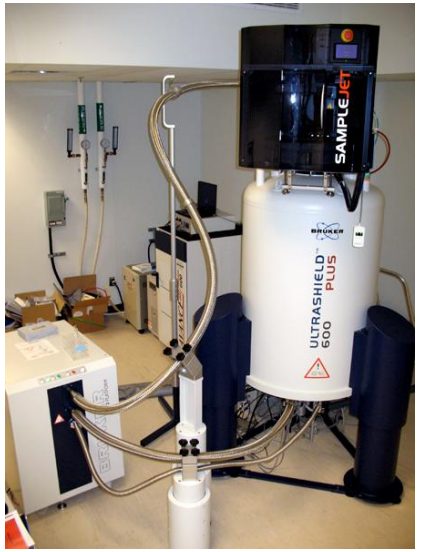
CD Crohn's disease; *EIM* extraintestinal manifestations; *HB* Harvey-Bradshaw score; *UC* ulcerative colitis; *P* proctitis; *PS* proctosigmoiditis; *LC* left-sided colitis; *PC* pancolitis; *IR* Ileocaecal resection; *HC* hemicolectomia; *C* colectomia

[a] Dose of 5 mg/day

Sample collection and preparation



^{1}H NMR spectroscopy



- NMR samples were prepared by mixing 4 µl of D2O/TSP (3-trimethylsilyl-2,2,3,3-tetradeuterosodium propionate) with 40 µl of fecal extract
- Bruker 600 MHz equipped with inverse detection cryogenic probe
- Spectra were acquired using CPMG pulse sequence with echo time of 160 ms
- 256 scans were collected with a spectral width of 12 ppm

Data Analysis

- NMR Spectra were manually corrected for phase and baseline distortions using TOPSPIN and referenced to TSP signal at δ 0.0 ppm
- Spectral regions δ 0.5-9.0 binned with width of 0.004 ppm using Amix
- Water regions δ 4.4-5.2 were removed to avoid saturation
- δ 2.12-2.26 which included region of 5-aminosalicylic acid and its metabolite were removed (drug used to treat UC patients)
- Data was normalized to total sum of their spectral integrals

Data Analysis

- PCA to visualize data
  - Data scaled to Unit variance
  - Hotelling's T^2 test to check for abnormalities
  - 10 samples were removed

- PLS-DA and OPLS-DA to classify data
  - Seven-fold cross validation
  - CV-ANOVA
  - Permutation test (n=200)
  - ROC curve (True positive rate vs False positive rate)

Data Analysis - Validation

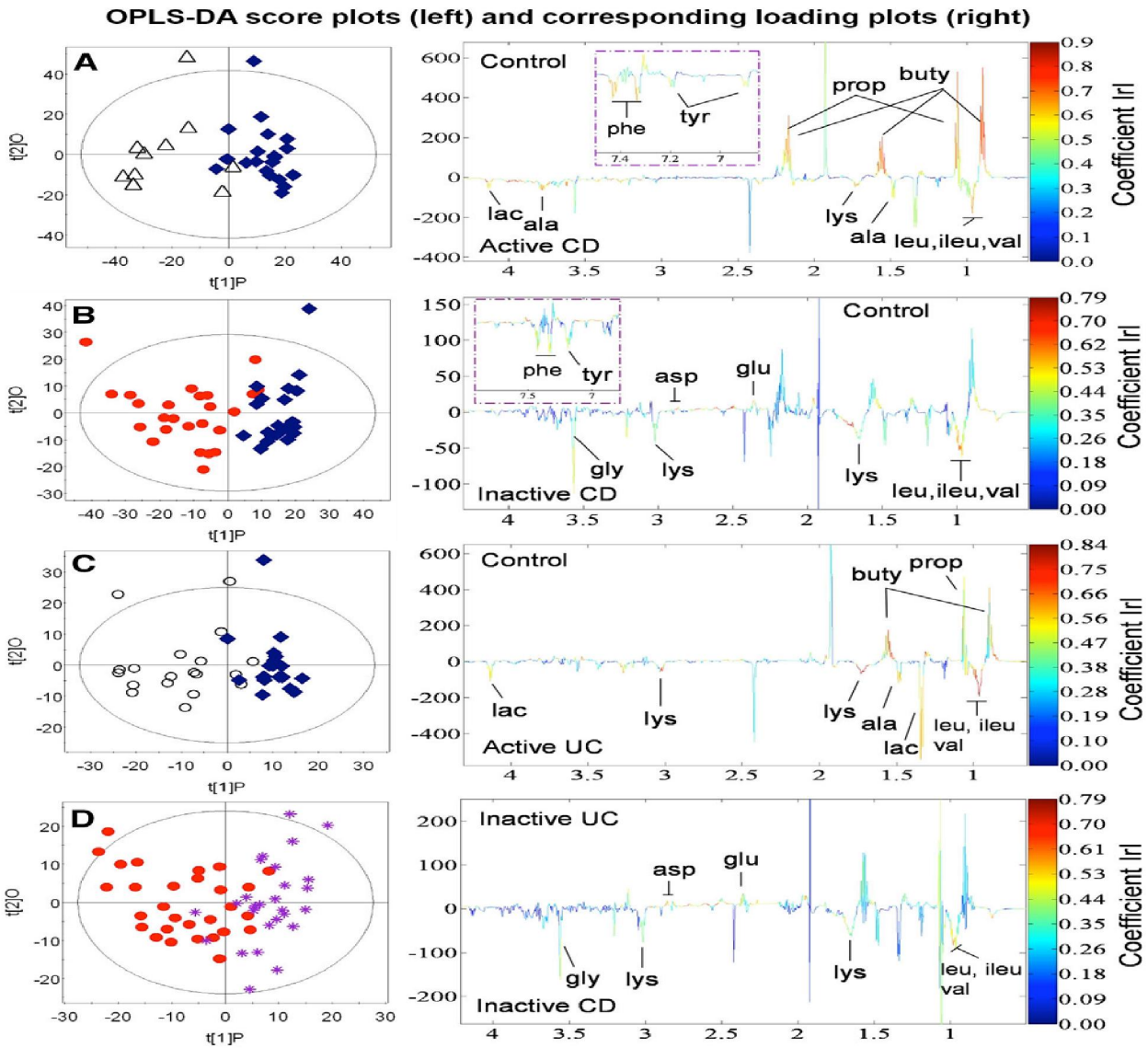
Table 2 Validation of PLS-DA and O-PLS-DA models

	CV-ANOVA (p-value)	Permutation (n = 200)
A) Models		
Active UC vs. inactive UC	0.0160 √	X
Active UC vs. controls	*0.0006* √	√
Inactive UC vs. controls	0.0297 √	X
Active UC vs. active CD	0.1083 X	X
Inactive UC vs. inactive CD	*0.0018* √	√
Active CD vs. inactive CD	0.2101 X	X
Active CD vs. controls	*0.0009* √	√
Inactive CD vs. controls	*0.0035* √	√
B) Models corrected for pouch/ileostomy		
Active UC vs. inactive UC	*0.0013* √	√
Active UC vs. controls	*0.0006* √	√
Inactive UC vs. controls	0.0065 √	X
Active UC vs. active CD	0.0973 X	X
Inactive UC vs. inactive CD	*0.0001* √	√
Active CD vs. inactive CD	1.0000 X	X
Active CD vs. controls	*0.0045* √	√
Inactive CD vs. controls	0.0191 √	X
C) Models corrected for intestinal surgery		
Active UC vs. inactive UC	*0.002* √	√
Active UC vs. controls	*0.000* √	√
Inactive UC vs. controls	0.013 √	X
Active UC vs. active CD	0.698 X	X
Inactive UC vs. inactive CD	0.317 X	X
Active CD vs. inactive CD	0.287 X	X
Active CD vs. controls	1.000 X	X
Inactive CD vs. controls	0.128 X	X

The models were only considered valid if the permutation test and the CV-ANOVA test ($p < 0.05$) were satisfied at the same time—italicized text

Results

Fig. 1 OPLS-DA score plots. The score plots (**a**, **b**, **c**, and **d**) are based on the four valid models containing all patients and display the 1st PLS component and one orthogonal component for each model. A two-way separation of the fecal samples is demonstrated in all 4 plots. *Blue diamonds* control; *empty triangles* active CD; *red dots* inactive CD; *empty circles* active UC; *purple stars*: inactive UC. The corresponding *back-scaled* loading plots reflect the class differences in the NMR spectra. *Upright peaks* indicate a relatively increased intensity of metabolites, and *downright peaks* a decreased intensity of metabolites. The *colors* shown on the plot are associated with the significance of metabolites in separating the samples as shown on the *right hand side* of the plot, where the *color-scaling map* is given together with the respective correlation coefficients. In accordance with the sample number in each group and a significance level of $p < 0.05$, the metabolites are significant at correlation coefficient values above **a** 0.55, **b** 0.43, **c** 0.44, and **d** 0.38, respectively. *CD* Crohn's disease; *UC* ulcerative colitis; *ala* alanine; *asp* aspartate; *buty* butyrate; *glu* glutamate; *gly* glycine; *ileu* isoleucine; *lac* lactate; *leu* leucine; *lys* lysine; *phe* phenylalanine; *prop* proprionate; *tyr* tyrosine; *val* valine (Color figure online)



Results

Table 3 Changed metabolites in the valid models

A)

Metabolites	Active CD/controls			Inactive CD/controls			Active UC/controls			Inactive CD/inactive UC		
	$Q^2 = 0.52$ $r > 0.55$	Student's *t* test (p-value)	ANOVA (p-value)	$Q^2 = 0.32$ $r > 0.43$	Student's *t* test (p-value)	ANOVA (p-value)	$Q^2 = 0.43$ $r > 0.44$	Student's *t* test (p-value)	ANOVA (p-value)	$Q^2 = 0.28$ $r > 0.38$	Student's *t* test (p-value)	ANOVA (p-value)
ileu	↑	0.002*	0.00*	↑	0.230	1.00	↑	0.009*	0.14	↑	0.307	1.00
leu	↑	0.001*	0.01*	↑	0.013*	0.12	↑	0.000*	0.00*	↑	0.075	0.54
val	↑	0.002*	0.02*	↑	0.014*	0.14	↑	0.001*	0.01*	↑	0.056	0.44
lys	↑	0.001*	0.01*	↑	0.092	0.62	↑	0.000*	0.00*	↑	0.113	0.70
ala	↑	0.001*	0.01*	–	0.294	0.94	↑	0.003*	0.03*	–	0.878	1.00
tyr	↑	0.044*	0.12	↑	0.074	0.97	–	–	–	–	–	–
phe	↑	0.003*	0.00*	↑	0.695	0.21	–	–	–	–	–	–
gly	↑	0.034*	0.29	↑	0.042*	0.35	–	0.016*	0.15	↑	0.111	0.69
buty	↓	0.000*	0.00*	–	0.372	0.99	↓	0.179	0.86	–	0.991	1.00
prop	↓	0.051*	0.40	–	0.201	0.89	↓	0.147	0.80	–	0.036*	0.31
lac	–	0.226	0.92	–	0.474	0.99	↑	0.044*	0.36	–	0.459	1.00
asp	–	0.064	0.57	↓	0.044*	0.36	–	0.708	0.99	↓	0.001*	0.01*
glu	–	0.307	1.00	↓	0.001*	0.06	–	0.222	1.00	↓	0.000*	0.00*

B)

Metabolites	Active UC/inactive UC			Active UC/controls		
	$Q^2 = 0.34$ $r > 0.44$	Student's *t* test (p-value)	ANOVA (p-value)	$Q^2 = 0.46$ $r > 0.44$	Student's *t* test (p-value)	ANOVA (p-value)
ileu	↑	0.016*	0.030*	↑	0.009*	0.019*
leu	↑	0.001*	0.000*	↑	0.000*	0.000*
val	↑	0.002*	0.000*	↑	0.001*	0.000*
lys	↑	0.000*	0.000*	↑	0.000*	0.000*
ala	↑	0.012*	0.007*	↑	0.003*	0.001*
buty	↓	0.526	1.000	↓	0.179	1.000
prop	↓	0.394	1.000	↓	0.147	0.946
lac	↑	0.050	0.012*	↑	0.044*	0.015*
tau	↑	0.014*	0.029*	↑	0.002*	0.002*

A) ↑ increased or ↓ decreased compared to control and inactive UC. Valid models based on all included samples

B)↑ increased or ↓ decreased compared to inactive UC and control. Valid models based on samples from patients, who have not had intestinal surgery

Ala alanine; *asp* aspartic acid; *buty* butyrate; *gly* glycine; *glu* glutamate; *ileu* isoleucine; *lac* lactate; *leu* leucine; *lys* lysine; *phe* phenylalanine; *prop* propionate; *tau* taurine; *tyr* tyrosine; *val* valine; *ANOVA* analysis of variance; *CD* Crohn's disease; *UC* ulcerative colitis; Q^2 predictability of the model; *r* correlation coefficient

* significance ($p < 0.05$)

Conclusions

- Study demonstrates both possibilities and limitations of the approach
- The models are able to classify with all the patients included
- Only two models work once patients with surgery are removed
- Surgery (even a minor one) has a drastic effect on fecal metabolome

Table 4 Predictive capability of the models

Models corrected for surgery	Area under the curve	
Active UC vs. inactive UC	0.813	Good
Active UC vs. controls	0.876	Good
Inactive UC vs. controls	0.654	Poor
Active UC vs. active CD	0.693	Poor
Inactive UC vs. inactive CD	0.659	Poor
Active CD vs. inactive CD	0.529	Poor
Active CD vs. controls	0.667	Poor
Inactive CD vs. controls	0.726	Fair

Prediction performance estimates presented as area under the curve (AUC) for each model