POLITECNICO DI TORINO
Repository ISTITUZIONALE

Assessment of dietary exposure related to dietary GI and fibre intake in a nutritional metabolomic study of human urine

Original
Assessment of dietary exposure related to dietary GI and fibre intake in a nutritional metabolomic study of human urine / Rasmussen, L. G.; Winning, H.; Savorani, F.; Ritz, C.; Engelsen, S. B.; Astrup, A.; Larsen, T. M.; Dragsted, L. O.. - In: GENES & NUTRITION. - ISSN 1555-8932. - ELETTRONICO. - 7:2(2012), pp. 281-293. [10.1007/s12263-011-0250-x]

Availability:
This version is available at: 11583/2815647 since: 2020-04-23T14:40:21Z

Publisher:
BMC

Published
DOI:10.1007/s12263-011-0250-x

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

(Article begins on next page)

12 October 2024

Genes Nutr (2012) 7:281–293
DOI 10.1007/s12263-011-0250-x

RESEARCH PAPER

Assessment of dietary exposure related to dietary GI and fibre intake in a nutritional metabolomic study of human urine

Lone G. Rasmussen · Hanne Winning · Francesco Savorani · Christian Ritz · Søren B. Engelsen · Arne Astrup · Thomas M. Larsen · Lars O. Dragsted

Received: 10 December 2010 / Accepted: 21 September 2011 / Published online: 9 October 2011
© Springer-Verlag 2011

Abstract There is a need for a tool to assess dietary intake related to the habitual dietary glycaemic index (GI) and fibre in groups with large numbers of individuals. Novel metabolite-profiling techniques may be a useful approach when applied to human urine. In a long-term, controlled dietary intervention study, metabolomics were applied to assess dietary patterns. A targeted approach was used to evaluate the effects on urinary C-peptide excretion caused by the dietary treatments. Seventy-seven overweight subjects followed an 8-week low-calorie diet (LCD) and were then randomly assigned to a high-GI or low-GI diet for 6 month during which they completed 24-h urine collections at baseline (prior to the 8-week LCD) and after randomisation to the dietary intervention, at month 1, 3 and 6, respectively. Metabolite profiling in 24-h urine was performed by ^{1}H NMR and chemometrics. Partial least squares (PLS) analysis indicated that urinary formate could discriminate between high-GI and low-GI diets (correlation coefficient $r = 0.82$), and this finding was confirmed statistically ($P = 0.01$). PLS analysis also indicated that urinary hippurate could be associated with fibre intake, but this finding was not confirmed statistically. No associations between GI and urinary C-peptide were found. Our results emphasise that application of metabolomics is useful in the assessment of dietary exposure related to dietary GI and fibre seen at group level in a nutritional metabolomic study of human urine. As our design allowed for large variations in individually selected food items, biomarkers identified at group level may be interpreted as more general and robust markers, largely not confounded with markers from single dietary factors.

Keywords Nutritional metabolomics · Glycaemic index · Dietary assessment · 1H NMR · Chemometrics · 24-h urine

Abbreviations

CTR	Control diet (name of intervention group)
DiOGenes	Diet, obesity and genes
GL	Glycaemic load
GLM	General linear models
HGI	High glycaemic index (name of intervention group)
HP	High protein (name of intervention group)
iPLS	Interval partial least squares regression
iPLS-DA	Interval discriminant partial least squares regression
LCD	Low-calorie diet
LGI	Low glycaemic index (name of intervention group)
LP	Low protein (name of intervention group)
LV	Latent variables
NMR	Nuclear magnetic resonance
PC	Principal component
PCA	Principal component analysis

L. G. Rasmussen (✉) · A. Astrup · T. M. Larsen · L. O. Dragsted
Department of Human Nutrition, Faculty of Life Sciences, University of Copenhagen, Frederiksberg, Denmark
e-mail: logr@life.ku.dk

H. Winning
Foss Analytical A/S, Hillerød, Denmark

F. Savorani · S. B. Engelsen
Quality and Technology, Department of Food Science, Faculty of Life Sciences, University of Copenhagen, Frederiksberg, Denmark

C. Ritz
Department of Basic Sciences and Environment, Faculty of Life Sciences, University of Copenhagen, Frederiksberg, Denmark

ppm	Parts per million
RMSECV	Root mean square error of cross-validation
TSP	Trimethylsilyl proprionic acid

Introduction

An increasing number of nutritional studies suggest that diets high in fibre and low in GI have a number of health benefits (Anderson et al. 1994; Rizkalla et al. 2002). Yet, most such studies have relied on subjective reporting of food diaries or food frequency questionnaires. However, one of the most difficult and challenging problems in nutritional studies entails the limitations of various dietary assessment methods in order to clarify the causal associations that require evidence of risk from exposure to particular dietary eating patterns. Therefore, there is a need for a tool to assess dietary patterns related to the habitual dietary intake related to dietary GI and fibre intake with groups of large numbers of individuals. The field of nutritional metabolomics might be of great value for nutritional studies in assessing dietary exposure (Fave et al. 2009; Jenab et al. 2009; Penn et al. 2010).

The use of biomarkers has the potential to objectively measure dietary intake and give a more accurate ranking of intake than using traditional dietary assessment methods (Bingham 2002), and some biomarkers may provide improved correlations between reduced risk of certain diseases and the intake of different foods. Biomarkers based on recovery of certain nutrients directly related to intake have been developed successfully but there is an urgent need to develop less targeted and more high-throughput methods with the opportunity to allow a more global characterisation of dietary patterns.

In short-term nutritional studies (Lenz et al. 2004; Solanky et al. 2003; Stella et al. 2006; Walsh et al. 2007; Wang et al. 2005), metabolomics have already shown promising results in linking metabolite contents in human biofluids to acute or chronic dietary exposure, but the application of metabolomics in long-term studies and when applied to free-living subjects remains to be investigated. The purpose of the present study was to apply a metabolomic approach for measuring influence of different dietary patterns related to ad libitum dietary intake on the urine metabolome in a 6-month randomised, controlled dietary intervention study with fixed goals for GI and fibre in each intervention group. The metabolomics analysis was based on samples from the Danish limb of the diet, obesity and genes (DiOGenes) study, which is a pan-European, multicenter, randomised, dietary intervention study. The DiOGenes intervention study was mainly designed with the objective to test the effects of high versus low protein and low- versus high-GI diets to maintain weight loss, when provided ad libitum to whole families with obese adults (Larsen et al. 2010a). However, the present study provides an opportunity to apply metabolomics to study controlled dietary patterns in a design that allows variation induced by the individual choices of food items thus mimicking the variation between free-living subjects

Subjects and methods

Study design

The present study is an extension of the work that was carried out in the Pan-European DiOGenes study. The main aspects of the research conducted in the comprehensive, long-term, randomised, controlled dietary intervention study were to address the impact of dietary protein and glycaemic index (GI) on weight (re)gain in a large number of families in which parents and children suffer from obesity or overweight. A detailed description of the study and further information on the study recruitment, exclusion criteria and the investigations carried out at the clinical investigation days have been described in two methodological papers (Larsen et al. 2010b; Moore et al. 2010). The study population in the present study consisted of the adult members of the families participating in the intervention study carried out in the Danish centre.

The study is registered at ClinicalTrials.gov, number NCT00390637.

Subjects

A total of 109 healthy, non-diabetic, obese subjects with a body mass index of 30.7–37.2 kg/m^2, aged 37–45 years underwent the clinical investigation day at baseline, which included collection of 24-h urine samples. All subjects then underwent an 8-weeks low-calorie diet treatment before starting the dietary interventions. Out of these, a total of 101 subjects collected 24-h urine samples after 1 month of intervention and 95 subjects collected 24-h urine samples after 3 month of the dietary intervention period. Finally, 77 (44 women, 33 men) of these subjects also collected 24-h urine samples and underwent the clinical investigation day at month 6. Metabolomic analyses of the current study are solely based on the 77 subjects, who completed the 24-h urine collections at all four time points. A schematic presentation of the study design is given in Fig. 1.

Experimental diets

The current study was designed as a parallel intervention trial with 5 dietary intervention groups, and the subjects

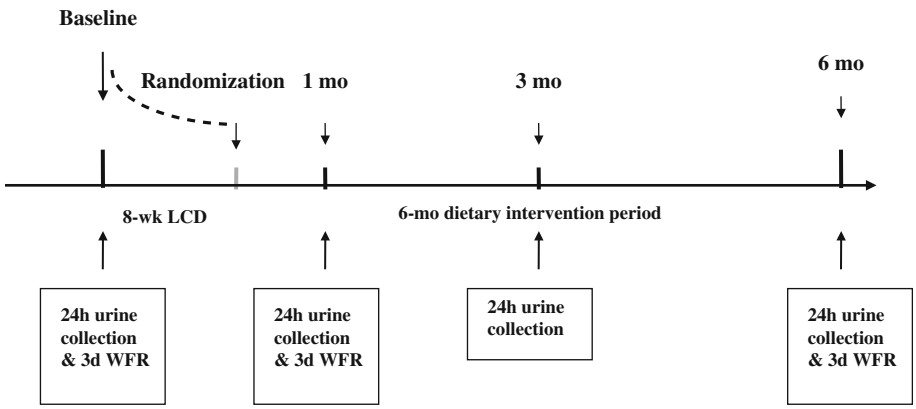

Fig. 1 Timeline of the part of the DiOGenes study included in the present study and the scheduled 24-h urine collections and 3-day weighed food records (WFR). *LCD* low-calorie diet

were randomly assigned to a 6-month low-fat (25–30% of energy) diet based on one of the following interventions: low GI, low protein (LGI/LP); low GI, high protein (LGI/HP); high GI, low protein (HGI/LP); high GI, high protein (HGI/HP) or to a control (CTR) diet according to the current Danish official guidelines reporting no specific recommendations for GI and an intermediate content of protein. The CTR diet group was therefore left out in the metabolomic analyses in the current paper due to the non-specific dietary GI.

The subjects were asked to complete a 3-day weighed food record (3-day WFR) for three consecutive days including two week days and one weekend day, both at baseline (prior to the 8-week LCD) and after randomisation to the dietary intervention, at month 1 and 6, respectively. The subjects were instructed to weigh all their foods whenever possible and to supply information on brand names, processing and cooking. When weighing was not possible (e.g. when dining out), they were instructed to record the food in household measures (cups, glasses, tablespoons, etc.). All foods noted in these diaries were coded to foods listed in a specific food database. By means of a standardised procedure combining the weight, the coding and the nutrient information for each food item, the nutrient intake was thereafter calculated for each food diary (Aston et al. 2010a).

To mimic free-living conditions, the 6-month dietary intervention was based on an ad libitum design where the families were provided with most food, free of charge, from a specially established trial supermarket at the university department. The validated supermarket model has been described elsewhere (Rasmussen et al. 2007).

A computer program was constructed for the recording of foods (product database) and for the calculation of nutrient composition of each shopping session during the 6-month dietary intervention (DiOGenes, version 1.4; Scientific Nutrition Supervision, Greve, Denmark). Local food manufacturers donated most of the products to the supermarket. Additional products were purchased to ensure an appropriate assortment to cover the dietary needs and the variability required by all diet groups throughout the 6-month period.

With respect to GI, the aim was to achieve a difference of 15 GI points in the LGI diets, compared with the HGI diets. The LGI/LP and HGI/LP diets were designed to provide 57–62% of energy from carbohydrate (CHO) and 10–15% of energy from protein (the energy percentage of fat remained constant), whereas the LGI/HP and HGI/HP diets prescribed 45–50% of energy from CHO and 23–28% of energy from protein. Thus, in the analyses where the two LGI diets were pooled, the diets had a range between 45–50% and 57–62% of energy from CHO and a range between 10–15% and 23–28% of energy from protein. The latter implies that both the LGI and the HGI diets had quite a substantial range in GL. A detailed description of the diets and the dietary strategy is described elsewhere (Aston et al. 2010b; Moore et al. 2010). Alcohol consumption was allowed in accordance with the current guidelines issued by the Danish National Board of Health, i.e., <14 units/week and <21 units/week (1 unit = 12 g alcohol) for women and men, respectively. Subjects were instructed to engage in minimum 30 min/day of moderate physical activity. All subjects were allowed a 3-week break from the project, during which no recording of the dietary intake was required. The 3-week break was randomly distributed throughout the 6 month among the study participants. However, the break was not allowed to be 1 week prior to the 24-h urine collections and the corresponding 3-day WFR.

24-h urine collection and storage

Twenty-four hour pooled urine samples were collected both at baseline and after 1, 3 and 6 month of the dietary intervention. Each subject was provided with two disinfected, pre-weighed and airtight 2.5-L polyethylene containers and a disinfected 500 mL container. The subjects were instructed not to collect the first urine void in the

morning on the first day, but after this first void, the urine collection continued until and including the morning urine void the following day. During the 24 h, the subjects were instructed to store the containers in the fridge or in another cool place, if possible (Rasmussen et al. 2010). The subjects were also instructed to take three para-aminobenzoic acid (PABA) tablets a day (240 mg/day), at the morning meal, at lunch and at dinner, respectively. This was to serve as a control of the completeness of urine collection, since PABA is absorbed and excreted in the urine within 24 h (Bingham 2002).

Urine sample handling

The total volume and the density of the collected 24-h urine were determined, and two 5 mL aliquots from each of the collections were drawn and stored at −80°C until further preparation and analysis, giving a storage period of between 1 and 16 month.

Urinary pro-insulin C-peptide was analysed by the IMMULITE 2500 C-peptide procedure and urinary PABA by spectrophotometry (Stasar, Gilford Instruments Laboratories, Oberlin, USA) (Bingham and Cummings 1983).

Prior to NMR analysis, the samples were thawed at 4°C, centrifuged at 4°C at 1,600 rpm for 10 min and then an 340 µL aliquot of the supernatant was added 170 µL 0.3 M sodium phosphate buffer (1 M perdeuterated 3-(trimethylsilyl) propionate sodium salt (TSP), 0.3% (w/v) NaN_3 and 60% (v/v) D_2O, pH 7.4). TSP was used for chemical shift reference, and NaN_3 was added as a preservative. A Gilson cooling rack kept the samples cool at 4°C prior to injection.

^{1}H NMR analysis

^{1}H NMR spectra were acquired on a Bruker DRX 600 MHz spectrometer (Bruker Biospin Gmbh, Rheinstetten, Germany) operating at 600,00 MHz for protons (14.09 Tesla) using a broadband inverse detection probe head equipped with a 120 µL flow cell. Data were accumulated at 300 K employing a pulse sequence composed by a pre-saturation of the water resonance during the recycle period followed by a composite 90° pulse with an acquisition time of 2.73 s, a recycle delay of 2 s, 128 scans and a sweep width of 12,019.23 Hz, resulting in 64 k complex data points. All samples were individually and automatically tuned, matched and shimmed. Prior to Fourier transformation, each 'free induction decay' (FID) was zero-filled to 64 k points and apodised by Lorentzian line broadening of 0.30 Hz. The resulting spectra were manually phased and automatically baseline corrected using Topspin™ (Bruker Biospin), and the ppm scale was referenced towards the TSP peak at 0.00 ppm.

Since the composition and the concentrations were very similar for every sample, the receiver gain was initially set at a fixed value equal to 700 in order to have a common intensity scale for all the acquired experiments.

Data pre-processing

The NMR spectra were corrected for signal misalignment due to the shift of small pH-dependent signals using the interval-based icoshift algorithm (Savorani et al. 2010a) and normalised according to the TSP signal for being more suitable for an effective alignment using icoshift. Only the spectral region between 9.20 and 0.62 ppm was considered, and the NMR region between 6.34 and 4.09 ppm was removed because NMR signals in this region (i.e. urea, α and β anomeric sugars) were strongly affected by the residual HDO peak.

Statistics and multivariate data analysis

Separate mixed linear models were used to assess the time–treatment interaction for the parameters carbohydrate and fibre intake, dietary GI and GL, and urinary C-peptide. The subjects individually subject numbers were included as random factors to account for heterogeneity between subjects. The analyses were based on the 77 subjects (LGI/LP: $n = 20$, LGI/HP: $n = 15$, HGI/LP: $n = 21$ and HGI/HP: $n = 21$) who completed the 24-h urine collections at baseline and after 1, 3 and 6 month. Post hoc pair-wise comparisons of treatments were made using t tests with Tukey–Kramer adjustment of the P values to maintain the pre-specified significance level and thus to minimise the risk of false positive findings. In case the interaction of time and treatment was significant, only the treatment differences for the final time were reported. It was tested whether gender, BMI, PABA, total urine volume, time, diet group and dietary components had a significant effect on the total urinary C-peptide excretion. Where appropriate, variables that significantly affected urinary C-peptide excretion were included as covariates in the analyses. Statistical analyses were performed with SAS (Statistical Analysis Package version 9.1 for Windows (SAS institute Inc, Cary, NC)), in particular the MIXED procedure. The significance level used was 0.05.

Out of 376 spectra, 112 samples were removed from the dataset because of either outlying behaviour or inadequate PABA recovery. Twenty-two urine samples were considered as outliers due to the presence of ethanol [1.17 ppm], paracetamol (acetaminophen) [2.18 ppm] and ibuprofen [0.91 ppm], which had been consumed by the subjects during the intervention period. Furthermore, the presence of acetate [1.91 ppm] indicated bacterial contamination (Rasmussen et al. 2010). Only subjects who had provided

urine collections with a PABA recovery of more than 75% ($n = 264$) and thus indicating good 24-h urine collections were included in the subsequent data analysis.

To remove the large inter-individual variation between the samples, the ^{1}H NMR spectra were mathematically averaged according to the design shown in Fig. 2a, where urine spectra from approximately 5 subjects comprise each average. For the classification of HGI/LGI diet groups, the 80 baseline samples were excluded from the model since no randomisation had taken place at this point. The 104 female and the 80 male spectra from visit 1, 3 and 6 month were averaged into the four diets groups LGI/LP, LGI/HP, HGI/LP and HGI/HP (without control) for each gender yielding 24 averaged spectra (Fig. 2b). In search for biomarkers associated with dietary fibre intake, 116 female urine and 84 male urine NMR spectra from baseline, month 1 and 6 were further averaged into HGI and LGI diets for each gender at each of the three time points yielding a total of 12 averaged spectra (Fig. 2c). Samples from month 3 were not included in the analysis, as no 3-day WFR were obtained at this time point.

Initially, unsupervised exploration of the data was carried out by means of principal components analysis (PCA) (Hotelling 1933). Supervised exploration of dietary parameters was performed using partial least squares regression (PLS) (Barker and Rayens 2003; Wold et al. 1983) and interval PLS (iPLS) which is an extension of PLS where local PLS models were calculated on a number of subintervals of the spectrum (Larsen et al. 2006). Furthermore, in order to classify the urine NMR spectra according to their diet groups, partial least squares discriminant analysis (PLS-DA) was applied by means of an interval-based method, which was applied to extract relevant information in different spectral intervals, but avoiding interference from other spectral regions (Nørgaard et al. 2000). Summarising, for categorical reference variables such as diet group, the method is referred to as interval PLS discriminant analysis (iPLS-DA), whereas for quantitative reference variables such as fibre intake, it is simply called interval PLS (iPLS). Interval-based chemometric models have previously proven to be powerful exploratory tools in providing knowledge about informative regions of the

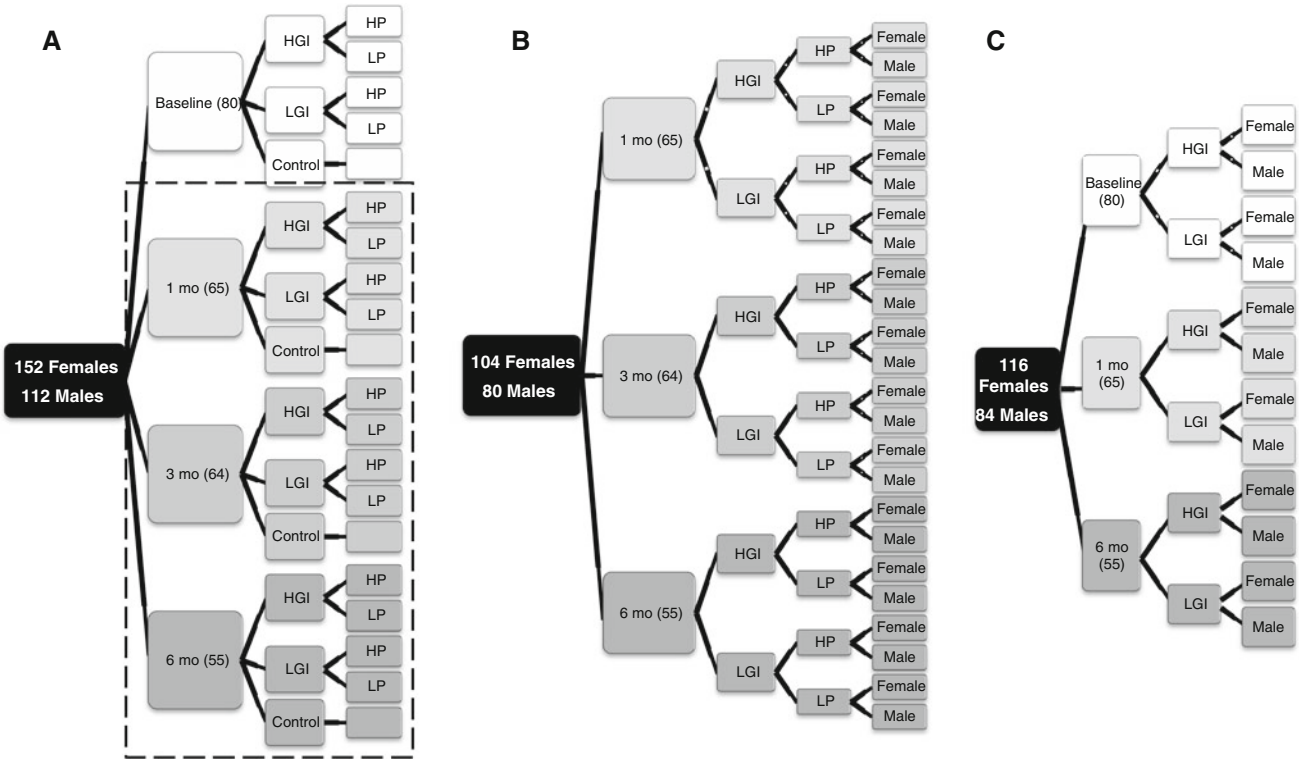

Fig. 2 **a** Averaging design, where urine spectra from approximately 5 subjects comprise each average. **b** For the classification of HGI/LGI diet groups, 80 samples from baseline were excluded from the model since no randomisation had taken place at this point. The 104 female and the 80 male spectra from visit 1, 3 and 6 month were averaged into the four diets groups LGI/LP, LGI/HP, HGI/LP and HGI/HP (without control) for each gender yielding 24 averaged spectra. **c** In the search for biomarkers associated with dietary fibre intake, 116 female and 84 male urine NMR spectra from baseline, month 1 and 6 were further averaged into HGI and LGI diets for each gender at each of the three time points yielding a total of 12 averaged spectra. Samples from month 3 were not included in the analysis, as no 3-day WFR were obtained at this time point

NMR spectrum (Kristensen et al. 2009; Rasmussen et al. 2010; Savorani et al. 2010b; Winning et al. 2009). For properly handling of this specific data, the intervals were determined such that each one contained 100 chemical shifts, yielding a total of 108 intervals. The number of latent variables (LV's) in the PLS model has been determined using cross-validation. The sample set is divided into a number of segments, which in turn are excluded 'one at a time' before re-entering into the model in order to estimate the prediction error, i.e., root mean square error of cross-validation (RMSECV). As a general rule, the optimal number of LV's is chosen from the first minimum of the RMSECV. Full cross-validation is used for the iPLS models where one average NMR spectrum is left out at a time, and the number of segments is the number of samples. In iPLS, RMSECV is calculated for each PLS model in each interval and compared to the RMSECV of the global model, which is based on the entire spectrum. Intervals with good discriminative power were identified by the lowest RMSECV. For the intervals identified as potentially separating diet groups, a confirmatory analysis of variance (ANOVA) based on the sums of maximum peak intensities was performed to compare diet groups statistically. The ANOVAs also included time (1, 3, 6 month) additively to adjust for possible time trends.

The data were analysed using the chemometric software Latentix 2.0. Matlab® (2009a, The Mathworks Inc., Natick, MA, USA) was used for both normalisation and signal alignment of spectral data. More specifically, the latter was performed using interval shifting by means of the previously mentioned *icoshift* algorithm (Savorani et al. 2010a). iPLS and iPLS-DA were also performed in Matlab using *iToolbox*. ANOVA was performed using SAS (Statistical Analysis Package version 9.1 for Windows (SAS institute Inc, Cary, NC)).

Results and discussion

Assessment of dietary intake and the effects on urinary C-peptide excretion—a targeted approach

Descriptive statistics and explorative PCA

The characteristics of the four diet groups are presented in Table 1. The dietary intake of carbohydrate and fibre complied with the stipulated diet in the four diet groups. Dietary GI was decreased in the LGI diets from the baseline measurement and to the measurements during the intervention and was found to be significantly lower in the two LGI diets compared to the HGI diets from month 1.

To provide an overview of the relationship between diet and the effect variables at the subject level at the end of the intervention period (month 6), a PCA model was calculated and the loading plot is shown in Fig. 3. The first two principal components (PC's) describe 50% of the variation. The first PC is describing variation in dietary energy with carbohydrate intake (starch (g/day), sugar (g/day), dietary fibre (g/day) and carbohydrate (as % of energy) and total dietary energy intake as the largest contributors in the positive direction of PC1 and protein (as % of energy) in the negative direction. The second PC reflects variation in dietary GI in the positive direction and total C-peptide excretion, Δ BMI, which is calculated as the BMI change between month 6 and 1 of the dietary intervention and fat (as % of energy) in the negative direction.

Statistical analysis of urinary C-peptide excretion and energy intake

A targeted approach was used to evaluate the effects on urinary C-peptide excretion caused by the dietary treatments. There was a significant main effect of time for total 24-h urinary C-peptide indicating that urinary C-peptide decreased from the baseline measurement and to the measurements during the intervention for all four diet groups ($P < 0.0001$). However, the difference in C-peptide excretion between the HGI and LGI diets was not statistically significant in this study and thus our results seem to corroborate previous findings. A study by Buyken et al. (2006) found that there was a significant association between urinary C-peptide of healthy children and total carbohydrate intake when adjusted for fibre intake, but for dietary GI, no clear association was found. Neither did Hartman et al. (2010) find any changes in fasting serum C-peptide in a randomised crossover study of 64 male participants consuming a legume-enriched, LGI diet versus a HGI healthy American diet. In contrast, previous studies did find that low-GI diets were able to lower urinary C-peptide (Jenkins et al. 1987, 1988; Wolever et al. 1992) in both healthy subjects and subjects with type 2 diabetes. In a study by Wu et al. (2004), the association of dietary fructose, GL and carbohydrate intake with fasting plasma C-peptide concentrations was studied and they found that high intakes of fructose and foods with a high GL were associated with higher C-peptide concentrations, whereas consumption of carbohydrates high in fibre was associated with lower C-peptide concentrations. However, no association was found between 24-h C-peptide excretion and dietary GI, GL, carbohydrate, sugar or fibre, respectively, as illustrated in Fig. 3 (based on univariate statistical analyses, not shown).

Table 1 Mean dietary intake of macronutrients/nutrients, GI and GL estimated by 3-day WFR and 24-h urinary C-peptide excretion during the 6-month diet intervention

	LGI/LP ($n = 20$)	HGI/LP ($n = 15$)	LGI/HP ($n = 21$)	HGI/HP ($n = 21$)	*P* Main effect of time	*P* Main effect of treatment	*P* Time × treatment
Men/women	8/12	4/11	11/10	10/11			
EI					<0.0001	NS	NS
Baseline (MJ)	8.95 ± 0.65[a]	11.96 ± 0.76	10.22 ± 0.62	10.16 ± 0.63			
Month 1 (MJ)	5.04 ± 0.65	6.17 ± 0.73	6.39 ± 0.62	7.36 ± 0.64			
Month 6 (MJ)	5.33 ± 0.63	5.36 ± 0.73	5.88 ± 0.66	5.37 ± 0.64			
CHO					–	–	<0.0001
Aim	57–62	57–62	45–50	45–50			
Baseline (% E)	46.2 ± 1.5	47.1 ± 1.9	44.7 ± 1.5	47.9 ± 1.5			
Month 1 (% E)	55.2 ± 1.7[b]	57.3 ± 1.7[c,d]	43.5 ± 1.5	48.7 ± 1.5			
Month 6 (% E)	52.3 ± 1.7	57.1 ± 1.8[e,f]	46.7 ± 1.7	48.4 ± 1.5			
Fibre					<0.0001	NS	NS
Baseline (g/day)	19.9 ± 2.5	20.7 ± 2.7	18.2 ± 2.3	19.5 ± 2.4			
Month 1 (g/day)	31.9 ± 2.6	26.1 ± 2.7	24.6 ± 2.4	21.7 ± 2.4			
Month 6 (g/day)	32.7 ± 2.7	25.8 ± 2.9	25.2 ± 2.7	22.0 ± 2.6			
Sugar					0.056	NS	NS
Baseline (g/day)	74.9 ± 6.2	62.2 ± 7.7	63.4 ± 5.9	66.3 ± 6.0			
Month 1 (g/day)	67.7 ± 7.0	92.1 ± 6.7	85.7 ± 6.1	80.8 ± 6.0			
Month 6 (g/day)	70.8 ± 7.2	84.5 ± 7.1	68.6 ± 6.9	80.8 ± 6.3			
Starch					0.014	NS	NS
Baseline (g/day)	104 ± 7.7	114 ± 9.7	108 ± 7.4	116 ± 7.6			
Month 1 (g/day)	140 ± 8.8	126 ± 8.3	124 ± 7.6	128 ± 7.5			
Month 6 (g/day)	136 ± 9.1	125 ± 9.1	136 ± 8.7	130 ± 8.0			
Protein					–	–	<0.0001
Aim	10–15	10–15	23–28	23–28			
Baseline (% E)	19.1 ± 0.9	17.9 ± 1.1	17.4 ± 0.9	16.8 ± 0.9			
Month 1 (% E)	15.0 ± 1.0	17.4 ± 0.9	23.8 ± 0.9	21.6 ± 0.9			
Month 6 (% E)	18.0 ± 1.0	17.7 ± 1.1	23.5 ± 1.0	25.1 ± 0.9			
Fat					<0.0001	NS	NS
Aim	25–30	25–30	25–30	25–30			
Baseline (% E)	31.9 ± 1.5	31.9 ± 1.9	33.3 ± 1.5	32.5 ± 1.5			
Month 1 (% E)	29.1 ± 1.7	26.2 ± 1.7	30.3 ± 1.5	28.4 ± 1.5			
Month 6 (% E)	27.6 ± 1.8	24.7 ± 1.8	26.3 ± 1.7	26.3 ± 1.6			
GI					–	–	<0.0001
Baseline (Units)	66.0 ± 1.2	64.2 ± 1.3	65.0 ± 1.1	63.3 ± 1.1			
Month 1 (Units)	59.0 ± 1.3[g]	63.2 ± 1.3	60.0 ± 1.1[h]	66.1 ± 1.2			
Month 6 (Units)	59.0 ± 1.4	64.8 ± 1.4	58.6 ± 1.4	63.4 ± 1.3			
GL					<0.0001	NS	NS
Baseline (g/day)	172 ± 12.5	230.5 ± 14.4	175.3 ± 11.6	193.1 ± 11.8			
Month 1 (g/day)	131.1 ± 13.1	129.1 ± 14.1	97.1 ± 11.8	129.8 ± 12.2			
Month 6 (g/day)	158.6 ± 13.9	156.9 ± 15.0	124.3 ± 13.3	127.3 ± 12.8			
U-C-peptide[i] (nmol/24 h)					<0.0001	NS	NS
Baseline	33.9 ± 3.3	33.5 ± 3.7	35.1 ± 3.1	27.2 ± 2.9			
Month 1	16.4 ± 3.7	19.7 ± 3.7	17.6 ± 3.2	17.1 ± 3.2			
Month 3	17.9 ± 3.4	21.1 ± 3.5	15.0 ± 3.4	17.7 ± 3.3			
Month 6	24.6 ± 3.9	20.7 ± 4.1	26.0 ± 3.8	18.0 ± 3.4			

Table 1 continued

	LGI/LP ($n = 20$)	HGI/LP ($n = 15$)	LGI/HP ($n = 21$)	HGI/HP ($n = 21$)	*P* Main effect of time	*P* Main effect of treatment	*P* Time × treatment
BMI					<0.0001	NS	NS
Baseline (kg/m^2)	35.6 ± 1.7	32.5 ± 2.1	34.8 ± 1.7	33.0 ± 1.8			
Month 1 (kg/m^2)	29.4 ± 2.0	25.4 ± 1.9	29.3 ± 1.8	26.7 ± 1.7			
Month 3 (kg/m^2)	31.3 ± 2.0	26.6 ± 2.0	29.7 ± 1.8	26.3 ± 1.7			
Month 6 (kg/m^2)	32.0 ± 2.0	28.8 ± 2.0	31.8 ± 1.9	29.4 ± 1.8			

LGI low glycaemic index, *HGI* high glycaemic index, *LP* low protein, *HP* high protein, *CHO* carbohydrate, *% E %* of energy

[a] x ± SE (all such values)

[b] Time × treatment: interaction $P < 0.0001$ as compared to LGI/HP, month 1

[c] Time × treatment: $P < 0.0001$ as compared to LGI/HP, month 1

[d] Time × treatment: $P = 0.005$ as compared to HGI/HP, month 1

[e] Time × treatment: $P = 0.002$ as compared to LGI/HP, month 6

[f] Time × treatment: $P = 0.015$ as compared to HGI/HP, month 6

[g] Time × treatment: $P = 0.005$ as compared to HGI/HP, month 1

[h] Time × treatment: $P = 0.017$ as compared to HGI/HP, month 1

[i] Significant covariate: gender

Energy intake and effects on urinary C-peptide excretion

Total energy intake within all four diet groups (reported by 3-day WFR at baseline, month 1, 3 and 6) decreased compared to the energy intake at baseline prior to the LCD. BMI of the subjects did not differ significantly between the four groups, whereas the main effect of time ($P < 0.0001$) showed that the subjects had achieved a significantly lower BMI after the LCD and before entering the 6-month dietary intervention period, which was expected and intentional. In our study, the energy intake during the 6-month dietary intervention was ad libitum in order to investigate the impact of the diets on weight maintenance after the low-calorie diet. However, the 3-day dietary recordings showed a significant decreased energy intake during the intervention partly because the subjects have been used to the restricted calorie intake during the LCD or simply caused by energy under-reporting, which is a common issue in self-reported dietary assessment methods (Bingham 2002; Bingham et al. 1997; Leiba et al. 2005; Schoeller 1990; Weber et al. 2001). Still, the markedly lower energy intake during the intervention in combination with the LCD-induced catabolism of body stores of fat into energy is expected to cause a major decrease in insulin secretion (Goyenechea et al. 2008). As the C-peptide level is decreased during weight loss, C-peptide excretion might reflect diet to a lesser extent than it would in a weight stable period.

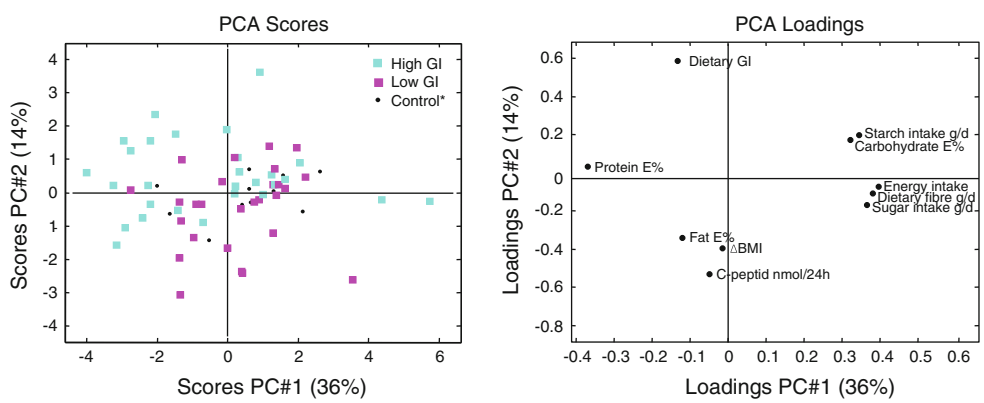

Fig. 3 PCA loading plot of the dietary variables (total energy intake (MJ/day), carbohydrate (% of energy), protein (% of energy), fat (% of energy), dietary fibre, sugar and starch intake (g/day) and dietary GI (units)) and the effect variable C-peptide (g/24 h) at month 6. ΔBMI is calculated as the BMI change between month 6 and 1. The loading plot illustrates the contribution of each variable to each PC. The first PC explains about 36% of the variation, and the second PC explains about 14% of the variation

Assessment of dietary exposure—a metabolomic approach

Pre-processing of the dataset included the removal of outliers due to bacterial appearance (acetate) in the urine or due to the intake of alcohol or medication by the subjects (presence of paracetamol (acetaminophen) [2.18 ppm] and ibuprofen [0.91 ppm] in the urine samples).

A PCA model of all the individual ^{1}H NMR spectra revealed large intra- and inter-individual variation and no groupings could be observed according to the dietary interventions (figure not shown). In order to reduce this variation, the spectra were averaged according to the design described in the *multivariate data analysis* section. This dataset consisting of averaged NMR spectra was used for the classification of the HGI and LGI samples in the following data analyses.

Associations between dietary GI and human urinary metabolites

Classification of the HGI and LGI diet groups by their urine spectral profile was performed using iPLS-DA calculated on 24 averaged NMR spectra (according to Fig. 2b). One outlying average spectrum was removed from the calculation. The following iPLS-DA model calculated with 2 latent variables (LV's) and 108 equally sized intervals revealed one interval that could improve the global model in the number of misclassification (Fig. 4). The corresponding actual versus predicted plot shows the model performance (Fig. 4): The correlation coefficient is 0.82 and a prediction error of 0.29. The model achieves a clear separation between the HGI and the LGI diets using the selected interval 8.43–8.49 ppm, and this finding was confirmed statistically ($P = 0.01$). The principal metabolite in this interval is found around 8.46 ppm and has been identified as formate (Wishart et al. 2005), which is likely responsible for the discrimination. The superimposition of urine NMR spectra of the signal from formate shows that the discrimination seems to be caused by higher average excretion of formate in the HGI diet groups than in the LGI diet groups (see small insert in Fig. 4). At 7.2 ppm, another signal also shows a low number of misclassifications (2/23 or 9%). To our knowledge, the activity seen in this interval is caused by yet unidentified metabolites.

Endogenous formate is a one-carbon product of fermentation by the gut microbiome where microbial reduction in CO_2 leads to the formation of formate (Kane and Breznak 1991). Kane et al. investigated how the host diet affected the production of organic acids and methane by the gut bacteria of cockroaches and found that the formation of formate and acetate from CO_2 was favoured after consuming a low-fibre feed compared to a high-fibre feed. Yet, Holmes et al. (2008) reported the finding of urinary formate excretion to be positively associated with energy intake; however, this was not the case in our study. The finding of formate in the present study therefore suggests that formate is excreted as a secondary metabolite by the HGI diets, which could be due to the significantly lower fibre intake found for these diets.

Associations between dietary fibre and human urinary metabolites

A metabolomic approach was applied to screen for biochemical markers at the group level associating with

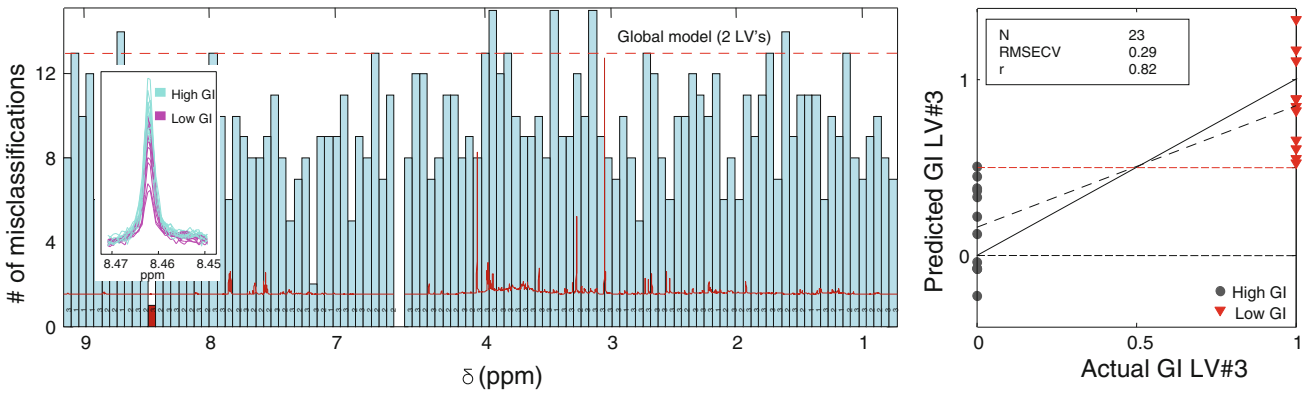

Fig. 4 *Left plot* Interval PLS-DA (iPLS-DA) plot of classification of the 4 diet groups (LP/LGI, LP/HGI, HP/LGI and HP/HGI) from 23 averaged ^{1}H NMR spectra of urine calculated on 108 intervals (full CV) showing the number of misclassification in absolute numbers together with the number of LV's. The *dashed red horizontal line* indicates the number of misclassification for the full spectral model using 2 LV's, and the *solid red line* is the average NMR spectrum. The small insert shows the 23 average signals from the metabolite around 8.46 ppm, which is linked to the metabolite formate. *Right plot* The actual versus predicted plot shows almost perfect separation of the HGI diets versus LGI diets using the selected interval. The RMSECV of the global model calculated with 3 LV′s is markedly improved for the interval in *red* (8.43–8.49 ppm). The *solid line* is reference for full agreement, whereas the *dotted line* is the best-fitting linear regression line

dietary fibre intake. Based on the 12 averaged urine NMR spectra divided into pooled HGI and LGI diets at baseline, month 1 and 6, it was possible to develop a good iPLS model of dietary fibre intake assessed by the 3-day WFR as shown in Fig. 5. The model was calculated on 12 averaged spectra from baseline, month 1 and 6 divided into HGI and LGI diet groups and gender (Fig. 2c). The iPLS plot indicates that the intervals around 3.96–4.02 and 7.57–7.70 ppm give the lowest prediction error, and the metabolite found to give rise to signals in this region was assigned as hippurate (Wishart et al. 2005). However, the PLS analysis also indicated that urinary hippurate could be associated with fibre intake, but this finding was not confirmed statistically. The predicted versus measured plot from this interval (illustrated in Fig. 5, right) shows that the correlation coefficient of the model was high ($r = 0.86$).

The results of this multivariate analysis suggest that hippurate is associated with dietary fibre intake at a group level in the studied population. Hippurate is known as a marker of increased consumption of polyphenol-rich foods like fruits and vegetables, nuts, grains, soy products and beverages such as red wine, tea and cocoa (Manach et al. 2004; Manach et al. 2005) or as a marker of consumption of foods containing benzoic acid, which may be present naturally or added as a preservative (Walsh et al. 2007). Wholegrain, fruits and vegetables are foods rich in dietary fibre, and the findings in the present study suggest that an increased consumption of fibre-rich foods on average also leads to the consumption of foods with a high content of polyphenols. Our findings support previous studies reporting that 24-h hippurate excretion was associated with dietary fibre (Holmes et al. 2008) and with a wholegrain diet in a pig study (Bertram et al. 2006).

Hippurate has also been reported as a metabolic end product of plant phenols and flavonoids by the gut microflora (Daykin et al. 2005; Mulder et al. 2005; Nicholson et al. 2005; Van Dorsten et al. 2006; Wang et al. 2005) and hence, it can also be speculated that hippurate is partially a metabolic signature by the gut microflora of the subjects participating in the present study.

In the current study, we did not observe other urinary profiles in the NMR spectra to be strongly related with fibre intake. However, alkylresorcinol measured in plasma and alkylresorcinol metabolites measured in urine have previously been reported as potential biomarkers of both wholegrain intake (Aubertin-Leheudre et al. 2008) and wholegrain rye and wheat cereal fibre (Aubertin-Leheudre et al. 2010; Bertram et al. 2006; Landberg et al. 2006, 2008, 2009). Common for all studies were that plasma and urine alkylresorcinol concentration analyses were performed by GC-MS or HPLC. Due to the reason that NMR is a far less sensitive analytical method, this may explain why we did not observe any signals originating from alkylresorcinols. Besides, alkylresorcinols are a marker of wholegrain rye and wheat, and another reason why we did not find any differences between groups could be due to the fact that we did not specifically control the intake of dietary wholegrain between groups.

Methodological considerations

A major strength of the present study design is the use of the supermarket model, which ensures a high degree of compliance to the diets. The supermarket system is assumed to be state-of-the-art method for controlling and monitoring dietary intake under ad libitum, free-living

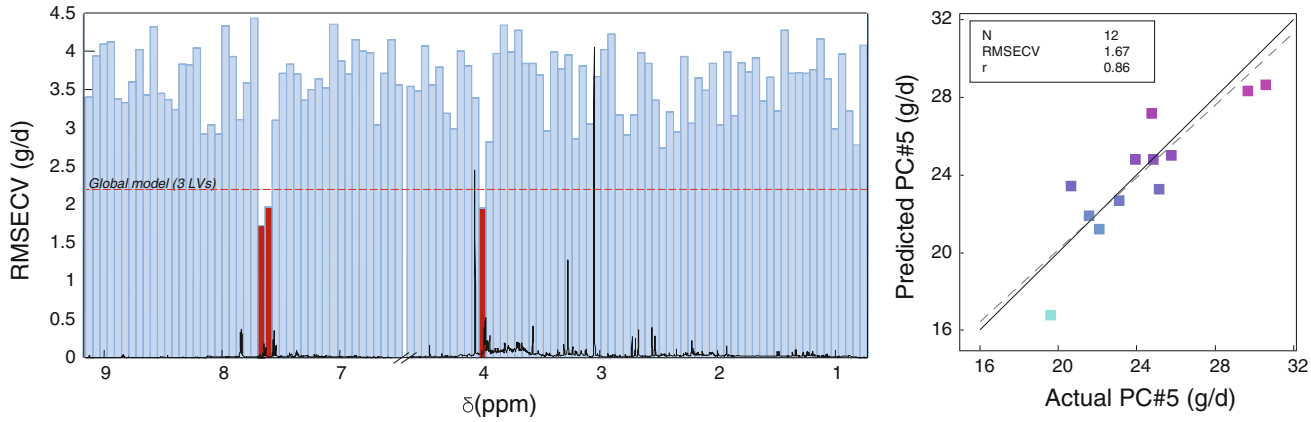

Fig. 5 *Left plot* Interval PLS (iPLS) plot calculated on 108 intervals (full CV) predicting fibre intake (g/day) (3-day WFR) from 12 averaged ^{1}H NMR urine spectra. The RMSECV of the global model calculated with 3 LV's is indicated by the *dashed red horizontal line*. The *red*-coloured intervals 3.96–4.02 and 7.57–7.70 ppm fall below the global RMSECV and have the smallest prediction errors. *Right plot* The predicted versus measured plot shows a RMSECV of 1.67 g/day and a correlation coefficient of 0.86. The *dots* are coloured according to the dietary fibre intake (*light blue* shows the lowest intake and *pink* shows the highest intake). The *solid line* is reference for full agreement, whereas the *dotted line* is the best-fitting linear regression line

conditions, although some uncertainty remains (Skov et al. 1997). It should be emphasised that despite the strict control of dietary intake regarding a number of macronutrients (i.e. protein, carbohydrate and fat), the design also allows for a quite substantial dietary variation as well as many other types of variation related to differences in lifestyle, intake of medication even within diet group. Hence, despite reflecting a more real-life situation, this in turn, also allows for a larger within-group variation. Also, there is of course no guarantee that the subjects actually consume the foods they selected in the supermarket during the 6-month diet intervention. The experience gathered in this study may be very useful for suggesting improvements to the supermarket model.

First of all, the design of the supermarket intervention in the DiOGenes study was family based with families consisting of a single adult with children, a couple with children or an adult participant with a non-participating partner (meaning that this person was considered non-eligible according to the inclusion criteria, but was still provided with food from the supermarket) and children. This type of intervention led to a shared household in the families, and therefore, the shopping sessions in the supermarket reflected the amount of food items provided to the whole family. Hence, the use of the supermarket database for estimating the dietary intake of each individual adult participant in the present study was considered less accurate than weighed food diaries. An improvement to the supermarket design would have included making the shopping sessions individually based. In addition, every shopping session recorded in the supermarket programme varied in the number of days until next session, and thereby, assessment of the individual dietary intake on a daily basis (and during the day) was not possible. The purpose of the DiOGenes study was not to measure dietary intake on a daily basis, and thus this highly controlled dietary intervention was perfectly suited for the estimation of dietary patterns on group level during the 6-month period.

As a result of the large variation in the individually selected food items, these biomarkers were identified at group level by using an average design (of the complete 376 urine NMR spectra) consisting of 12–40 averaged spectra divided into diet group and gender at baseline, month 1, 3 and 6. Obviously, results based on averaged urine NMR spectra are not as strong as results based on individuals, and this might limit the use of metabolomics-based biomarkers as this way of handling the data represents a huge simplification of the metabolomic data matrix, and meaningful intra-individual variation in response to the dietary intake is lost. On the contrary, this averaging design turned out to be a highly effective way of handling data with large variable patterns, and for this reason, larger sample sets may not overcome the problem concerning these large individual variations. Moreover, this design is closer to the 'real-life' situations for which we aim to apply metabolomics and the results therefore also point to major limitations in metabolomics and specifically to the use of analytical techniques with relatively low sensitivity, such as NMR, to create the dataset for analyses of highly variable patterns. It is possible that GC-MS or LC-MS analysis would reveal more details and thereby provide better data for metabolomics analyses of dietary patterns; however, the variability of these techniques made the choice of NMR a natural starting point. Therefore, an improvement to the present metabolomics study design would include additional analyses conducted by using more sensitive analytical techniques.

Concluding remarks

In the present study, we applied a holistic metabolomics approach where no a priori selection of metabolites was made. To our knowledge, this was the first long-term controlled human intervention study applying an exploratory approach to reveal the dietary exposure related to dietary GI and fibre intake in groups of large numbers of individuals participating in a long-term, randomised, controlled dietary intervention study. The supermarket model is a very interesting and potentially power way to guide food consumption in a manner that can emphasise certain food groups while maintaining the free-living features of the study, and this is the first study to incorporate this model in a global metabolomics analysis. By giving the participants a free choice of foods while controlling that they complied with the dietary patterns, this study strongly reflects a real-life situation, and the analyses are not confounded by markers relating to single individual food items, except total fibre and should only reveal markers of dietary GI and fibre intake. However, the issue with this model is that these controls appear to be too weak in this case to find much in the way of meaningful biomarkers related to the dietary groups.

In conclusion, our results emphasise that application of metabolomics in assessing dietary patterns controlled at the group level can be used to discover multivariate associations between dietary patterns and the excreted urinary metabolites in long-term dietary interventions.

Acknowledgments The authors responsibilities were as follows: AA, TML and LGR were responsible for the study design. LGR and TML were responsible for conducting the trial. FS was responsible for the ^{1}H NMR analyses. HW, SBE and LOD conducted the chemometric data analyses and LGR and CR conducted the statistics. SBE, HW, LOD and LGR were responsible for the interpretation of the results. LGR authored the manuscript with co-authorship from HW, FS, CR, SBE, AA, TML and LOD. All authors have read and approved the final manuscript. This project (the DiOGenes project) has been funded by a grant from the EU Food Quality and Safety

Priority of the Sixth Framework Programme, contract no. FP6-5 2005-513946 and was also supported by a grant from the Danish Ministry of Food, Agriculture and Fisheries (3304-FVFP-060706-01). A number of foods for the supermarket was provided free of charge from food manufacturers. A full list of these sponsors can be seen at http://www.diogenes-eu.org/sponsors/.

Conflict of interest There were no conflicts of interest for any of the authors.

References

Anderson JW, Smith BM, Gustafson NJ (1994) Health benefits and practical aspects of high-fiber diets. Am J Clin Nutr 59:1242S–1247S

Aston LM, Laccetti R, Mander AP et al (2010a) No difference in the 24-hour interstitial fluid glucose profile with modulations to the glycemic index of the diet. Nutrition 26:290–295

Aston LM, Jackson D, Monsheimer S et al (2010b) Developing a methodology for assigning glycaemic index values to foods consumed across Europe. Obes Rev 11:92–100

Aubertin-Leheudre M, Koskela A, Marjamaa A, Adlercreutz H (2008) Plasma alkylresorcinols and urinary alkylresorcinol metabolites as biomarkers of cereal fiber intake in Finnish women. Cancer Epidemiol Biomarkers Prev 17:2244–2248

Aubertin-Leheudre M, Koskela A, Samaletdin A, Adlercreutz H (2010) Plasma alkylresorcinol metabolites as potential biomarkers of whole-grain wheat and rye cereal fibre intakes in women. Br J Nutr 103:339–343

Barker M, Rayens W (2003) Partial least squares for discrimination. J Chemom 17:166–173

Bertram HC, Bach Knudsen KE, Serena A et al (2006) NMR-based metabonomic studies reveal changes in the biochemical profile of plasma and urine from pigs fed high-fibre rye bread. Br J Nutr 95:955–962

Bingham SA (2002) Biomarkers in nutritional epidemiology. Public Health Nutr 5:821–827

Bingham S, Cummings JH (1983) The use of 4-aminobenzoic acid as a marker to validate the completeness of 24 h urine collections in man. Clin Sci (Lond) 64:629–635

Bingham SA, Gill C, Welch A et al (1997) Validation of dietary assessment methods in the UK arm of EPIC using weighed records, and 24-hour urinary nitrogen and potassium and serum vitamin C and carotenoids as biomarkers. Int J Epidemiol 26(Suppl 1):S137–S151

Buyken AE, Kellerhoff Y, Hahn S, Kroke A, Remer T (2006) Urinary C-peptide excretion in free-living healthy children is related to dietary carbohydrate intake but not to the dietary glycemic index. J Nutr 136:1828–1833

Daykin CA, Van Duynhoven JP, Groenewegen A et al (2005) Nuclear magnetic resonance spectroscopic based studies of the metabolism of black tea polyphenols in humans. J Agric Food Chem 53:1428–1434

Fave G, Beckmann ME, Draper JH, Mathers JC (2009) Measurement of dietary exposure: a challenging problem which may be overcome thanks to metabolomics? Genes Nutr 4:135–141

Goyenechea E, Crujeiras AB, Abete I, Parra D, Martinez JA (2008) Enhanced short-term improvement of insulin response to a low-caloric diet in obese carriers the Gly482Ser variant of the PGC-1alpha gene. Diabetes Res Clin Pract 82:190–196

Hartman TJ, Albert PS, Zhang Z et al (2010) Consumption of a legume-enriched, low-glycemic index diet is associated with biomarkers of insulin resistance and inflammation among men at risk for colorectal cancer. J Nutr 140:60–67

Holmes E, Loo RL, Stamler J et al (2008) Human metabolic phenotype diversity and its association with diet and blood pressure. Nature 453:396–400

Hotelling H (1933) Analysis of a complex of statistical variables into principal components. J Educ Psychol 24:498–520

Jenab M, Slimani N, Bictash M, Ferrari P, Bingham SA (2009) Biomarkers in nutritional epidemiology: applications, needs and new horizons. Hum Genet 125:507–525

Jenkins DJ, Wolever TM, Collier GR et al (1987) Metabolic effects of a low-glycemic-index diet. Am J Clin Nutr 46:968–975

Jenkins DJ, Wolever TM, Buckley G et al (1988) Low-glycemic-index starchy foods in the diabetic diet. Am J Clin Nutr 48:248–254

Kane MD, Breznak JA (1991) Effect of host diet on production of organic acids and methane by cockroach gut bacteria. Appl Environ Microbiol 57:2628–2634

Kristensen M, Savorani F, Ravn-Haren G et al (2009) NMR and interval PLS as reliable methods for determination of cholesterol in rodent lipoprotein fractions. Metabolomics 129–136

Landberg R, Linko AM, Kamal-Eldin A et al (2006) Human plasma kinetics and relative bioavailability of alkylresorcinols after intake of rye bran. J Nutr 136:2760–2765

Landberg R, Kamal-Eldin A, Andersson A, Vessby B, Aman P (2008) Alkylresorcinols as biomarkers of whole-grain wheat and rye intake: plasma concentration and intake estimated from dietary records. Am J Clin Nutr 87:832–838

Landberg R, Aman P, Friberg LE et al (2009) Dose response of whole-grain biomarkers: alkylresorcinols in human plasma and their metabolites in urine in relation to intake. Am J Clin Nutr 89:290–296

Larsen FH, van den Berg F, Engelsen SB (2006) An exploratory chemometric study of H-1 NMR spectra of table wines. J Chemom 20:198–208

Larsen TM, Dalskov SM, van Baak M et al (2010a) Diets with high or low protein content and glycemic index for weight-loss maintenance. N Engl J Med 363:2102–2113

Larsen TM, Dalskov S, van Baak M et al (2010b) The diet, obesity and genes (Diogenes) dietary study in eight European countries—a comprehensive design for long-term intervention. Obes Rev 11:76–91

Leiba A, Vald A, Peleg E, Shamiss A, Grossman E (2005) Does dietary recall adequately assess sodium, potassium, and calcium intake in hypertensive patients? Nutrition 21:462–466

Lenz EM, Bright J, Wilson ID et al (2004) Metabonomics, dietary influences and cultural differences: a 1H NMR-based study of urine samples obtained from healthy British and Swedish subjects. J Pharm Biomed Anal 36:841–849

Manach C, Scalbert A, Morand C, Remesy C, Jimenez L (2004) Polyphenols: food sources and bioavailability. Am J Clin Nutr 79:727–747

Manach C, Williamson G, Morand C, Scalbert A, Remesy C (2005) Bioavailability and bioefficacy of polyphenols in humans. I. Review of 97 bioavailability studies. Am J Clin Nutr 81:230S–242S

Moore CS, Lindroos AK, Kreutzer M et al (2010) Dietary strategy to manipulate ad libitum macronutrient intake, and glycaemic index, across eight European countries in the Diogenes Study. Obes Rev 11:67–75

Mulder TP, Rietveld AG, van Amelsvoort JM (2005) Consumption of both black tea and green tea results in an increase in the excretion of hippuric acid into urine. Am J Clin Nutr 81:256S–260S

Nicholson JK, Holmes E, Wilson ID (2005) Gut microorganisms, mammalian metabolism and personalized health care. Nat Rev Microbiol 3:431–438

Nørgaard L, Saudland A, Wagner J et al (2000) Interval partial least squares regression (iPLS): a comparative chemometric study

with an example from near infrared spectroscopy. Appl Spectrosc 54:413–419

Penn L, Boeing H, Boushey CJ et al (2010) Assessment of dietary intake: NuGO symposium report. Genes Nutr 5:205–213

Rasmussen LG, Larsen TM, Mortensen PK, Due A, Astrup A (2007) Effect on 24-h energy expenditure of a moderate-fat diet high in monounsaturated fatty acids compared with that of a low-fat, carbohydrate-rich diet: a 6-mo controlled dietary intervention trial. Am J Clin Nutr 85:1014–1022

Rasmussen LG, Savorani F, Larsen TM et al (2010) Standardization of factors that influence human urine metabolomics. Metabolomics

Rizkalla SW, Bellisle F, Slama G (2002) Health benefits of low glycaemic index foods, such as pulses, in diabetic patients and healthy individuals. Br J Nutr 88(Suppl 3):S255–S262

Savorani F, Kristensen M, Larsen FH, Astrup A, Engelsen SB (2010a) High throughput prediction of chylomicron triglycerides in human plasma by nuclear magnetic resonance and chemometrics. Nutr Metab (Lond) 7:43

Savorani F, Tomasi G, Engelsen SB (2010b) icoshift: a versatile tool for the rapid alignment of 1D NMR spectra. J Magn Reson 202:190–202

Schoeller DA (1990) How accurate is self-reported dietary energy intake? Nutr Rev 48:373–379

Skov AR, Toubro S, Raben A, Astrup A (1997) A method to achieve control of dietary macronutrient composition in ad libitum diets consumed by free-living subjects. Eur J Clin Nutr 51:667–672

Solanky KS, Bailey NJ, Beckwith-Hall BM et al (2003) Application of biofluid 1H nuclear magnetic resonance-based metabonomic techniques for the analysis of the biochemical effects of dietary isoflavones on human plasma profile. Anal Biochem 323:197–204

Stella C, Beckwith-Hall B, Cloarec O et al (2006) Susceptibility of human metabolic phenotypes to dietary modulation. J Proteome Res 5:2780–2788

Van Dorsten FA, Daykin CA, Mulder TP, Van Duynhoven JP (2006) Metabonomics approach to determine metabolic differences between green tea and black tea consumption. J Agric Food Chem 54:6929–6938

Walsh MC, Brennan L, Pujos-Guillot E et al (2007) Influence of acute phytochemical intake on human urinary metabolomic profiles. Am J Clin Nutr 86:1687–1693

Wang Y, Tang H, Nicholson JK et al (2005) A metabonomic strategy for the detection of the metabolic effects of chamomile (*Matricaria recutita* L.) ingestion. J Agric Food Chem 53:191–196

Weber JL, Reid PM, Greaves KA et al (2001) Validity of self-reported energy intake in lean and obese young women, using two nutrient databases, compared with total energy expenditure assessed by doubly labeled water. Eur J Clin Nutr 55:940–950

Winning H, Roldan-Marin E, Dragsted LO et al (2009) An exploratory NMR nutri-metabonomic investigation reveals dimethyl sulfone as a dietary biomarker for onion intake. Analyst 134:2344–2351

Wishart DS, Knox C, Guo AC et al (2005) Human metabolome database. version 2.5. Retrieved Dec 2009

Wold S, Martens H, Wold H (1983) The multivariate calibration-problem in chemistry solved by the PLS method. Lect Notes Math 973:286–293

Wolever TM, Jenkins DJ, Vuksan V et al (1992) Beneficial effect of a low glycaemic index diet in type 2 diabetes. Diabet Med 9:451–458

Wu T, Giovannucci E, Pischon T et al (2004) Fructose, glycemic load, and quantity and quality of carbohydrate in relation to plasma C-peptide concentrations in US women. Am J Clin Nutr 80:1043–1049