Published in final edited form as:

Mol Nutr Food Res. 2015 April ; 59(4): 795–806. doi:10.1002/mnfr.201400847.

Identification and monitoring of metabolite markers of dry bean consumption in parallel human and mouse studies

Thushanthi Perera¹, Matthew R. Young², Zhiying Zhang³, Gwen Murphy⁴, Nancy H. Colburn², Elaine Lanza², Terryl J. Hartman⁵, Amanda J. Cross⁶, and Gerd Bobe¹

¹Linus Pauling Institute and Department of Animal and Rangeland Sciences, Oregon State University, Corvallis, OR 97331

²Laboratory of Cancer Prevention, Center for Cancer Research (CCR), National Cancer Institute (NCI), National Institutes of Health (NIH), Department of Health and Human Services (DHHS), The Pennsylvania State University, University Park, PA 16871

³Department of Nutritional Sciences, The Pennsylvania State University, University Park, PA 16871

⁴Nutritional Epidemiological Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute (NCI), National Institutes of Health (NIH), Department of Health and Human Services (DHHS), Rollins School of Public Health, Emory University, Atlanta, GA 30322

⁵Department of Epidemiology, Rollins School of Public Health, Emory University, Atlanta, GA 30322

⁶Department of Epidemiology and Biostatistics, Imperial College London, St Mary's Campus, London W2 1PG, U.K.

Abstract

Scope—Aim of the study was to identify and monitor metabolite markers of dry bean consumption in parallel human and mouse studies that each had shown chemopreventive effects of dry bean consumption on colorectal neoplasia risk.

Methods and Results—Using liquid chromatography/mass spectroscopy +/- electrospray ionization and gas chromatography/mass spectroscopy, serum metabolites of dry beans were measured in 46 men before and after a four-week dry bean-enriched diet (250 g/d) and 12 mice that received a standardized diet containing either 0 or 10% navy bean ethanol extract for 6 weeks; we also investigated fecal metabolites in the mice. The serum metabolites identified in these controlled feeding studies were then investigated in 212 polyp-free participants from the Polyp Prevention Trial who self-reported either increased (+31 g/d from baseline), high dry bean intake of 42 g/d in year 3 or low, unchanged dry bean consumption of <8 g/d; serum was analyzed from baseline and year 3. Serum pipercolic acid and S-methyl-cysteine were elevated after dry bean

Correspondence: Gerd Bobe, 112 Withycombe Hall, Oregon State University, Corvallis, OR97331; gerd.bobe@oregonstate.edu; phone: 541-737-1898; fax: 541-737-4174.

Authors contributions: Study design and experimental plan (MRT, EL, GM, NC, TJH, AJC, GB). Sample preparation (MRY, ZZ, TJH, GB). Sample and data analysis (TP, GB). Manuscript preparation (TP, MR, GM, TJH, AJC, GB).

Conflict of interest: None of the authors have a conflict of interest.

consumption in human and mouse studies and reflected dry bean consumption in the Polyp Prevention Trial.

Conclusions—Serum levels of pipecolic acid and S-methyl-cysteine are useful biomarkers of dry bean consumption.

Keywords

biomarkers; dry bean consumption; metabolomics; pipecolic acid; S-methyl-cysteine

1 Introduction

Colorectal cancer (CRC) is an important public health problem, accounting worldwide for over 694,000 deaths in 2012 [1]. In the U.S., CRC was the fourth most common cancer (135,260 new cases) and the second leading cause of cancer-related deaths (51,738 deaths) in 2011 [2]. Costs for CRC treatment were estimated to be over \$14 billion for 2010 in the U.S. [3]. Prevention strategies are imperative.

Diet is an established risk factor for CRC and presents a safe and effective strategy for CRC prevention [4]. Consumption of legumes, such as peas, beans, lentils, chickpeas and soybeans, is inversely associated with colorectal neoplasia risk in humans [5]. In a series of human and mouse intervention studies, we showed that high intake of dry beans, such as baked, kidney, pinto, lima, black, and navy beans, decreases colorectal neoplasia. In the 4-year Polyp Prevention Trial (PPT), increased self-reported dry bean consumption decreased serum markers of inflammation and insulin resistance and ultimately decreased advanced adenoma recurrence [6–8]. In the Legume Inflammation Feeding Experiment (LIFE), a controlled human feeding study, a 4-wk high dry bean diet (250 g/d) favorably changed serum markers of inflammation, insulin resistance, and hypercholesterolemia [9, 10], which are associated positively with CRC. In *ob/ob* mice developing azoxymethane (AOM)-induced tumors, we documented that the preventive effect of navy beans was strongest in the navy bean ethanol extract [11].

To demonstrate a protective effect of dry beans on colorectal neoplasia occurrence and progression in humans, biomarkers of dry bean intake are needed. These markers can monitor dietary compliance in intervention studies and reduce misclassification from self-reported dietary assessments, especially when dry beans are consumed as part of mixed dishes. Biomarkers in serum and fecal samples can measure not only recent intake of bioactive compounds but also estimate bioavailability at the tissue level.

Currently, there are very few established dietary biomarkers and none for dry beans. Metabolomics provides a global analysis of metabolites in biological samples and is a novel and promising tool for identifying exposure biomarkers in humans [12, 13]. Two essential criteria of a useful serum dry bean marker are that the biomarker is abundant in and relatively specific to dry beans and that the biomarker has a strong association with dry bean consumption or changes in dry bean intake, that are consistent across various populations. Thus, the first objective of our study was to identify metabolite markers that are unique to dry bean consumption using controlled human and mouse feeding studies. The second

objective was to evaluate whether the identified metabolite markers reflected dry bean consumption in samples of polyp-free subjects from a multi-year human intervention study that promoted their consumption. Two markers emerged, namely pipercolic acid and S-methyl-cysteine, that reflected dry bean consumption.

2 Materials and methods

2.1 Controlled human feeding study (LIFE)

The Legume Inflammation Feeding Experiment (LIFE) was a randomized controlled cross-over feeding study conducted at the General Clinical Research Center (University Park, PA) to evaluate whether a 4-week high dry bean diet (250 g or 1.5 cups/day of cooked pinto, navy, kidney, lima, and black beans) altered serum markers of insulin resistance or inflammation. All aspects of this study were approved by the Institutional Review Boards of the Pennsylvania State University (IRB21051) and the National Cancer Institute (05CN215-A) and are described in detail elsewhere [9]. Non-smoking males, aged 35 to 75 years, who had undergone colonoscopies within the previous 2 years, were randomly assigned to either a 4-week high dry bean or a healthy American diet (chicken meat replaced dry beans as primary protein source) and then were switched after a 2-week washout period to the other diet for the second 4-week feeding period. Diets were isocaloric and similar in macronutrients. Portion sizes were adjusted to avoid weight changes of participants. We selected an equal number of participants that were insulin resistant, had a colorectal adenoma history, had both, or had neither ($n = 12$ per category; samples for two individuals without insulin resistance and colorectal adenoma history were lost). We analyzed 12-hour overnight fasting serum samples collected at the beginning and end of the dry bean diet period. Samples were stored at -80°C until the end of study for analysis.

2.2 Self-reported dry bean consumption study (PPT)

The PPT, was a 4-year randomized, multi-center, nutritional intervention trial to evaluate whether a high-fiber (4.30 g/MJ or 18 g/1,000 kcal), high-fruit and high-vegetable (0.84 servings/MJ or 5 servings/d), and low-fat (20% of energy) diet is effective in inhibiting colorectal adenoma recurrence. The study was approved by the institutional review boards of the National Cancer Institute and those of the collaborating centers (OH91C0159-B). The study was registered with the [ClinicalTrials.gov](https://clinicaltrials.gov) identifier NCT00339625. A detailed description of the study has been published elsewhere [14, 15]. Briefly, men and women, aged 35 years or older, with at least one histologically confirmed colorectal adenoma removed in the prior 6 months, were randomized at baseline (T0) to the dietary intervention or control group for 4 consecutive years of follow-up. At T0 and at the end of each year (T1 to T4), participants provided a blood sample and completed an interviewer-administered questionnaire about demographic, clinical, medication and supplement use, and a food frequency questionnaire (FFQ) querying usual diet during the previous year. A single question was asked about the intake of cooked dry beans, such as pinto, navy beans, lentils and bean soups (amount and frequency). The 5 most commonly consumed dry beans based on annually administered four-day food records, available for a subset of participants, were in this order: baked, kidney, pinto, lima, and navy beans [6].

Using a nested study design, we selected participants (men and women) who were polyp-free at T4 with fasting blood samples available from T0 and T3, as well as complete FFQs from both of these time points. Participants (n=106 per group) self-reported either increased (+31 g/d T3-T0), high dry bean intake of 42 g/d (T3) or low, unchanged dry bean consumption of <8 g/d at T0 and T3. Individuals with and without increased dry bean consumption were matched on gender, age (+/- 5 years) and changes in consumption of other dietary components (i.e., fruits, vegetables, fat, flavonols, and fiber). In addition, we selected 21 participants, which did not fit into either category. We analyzed 12-hour fasting serum samples collected at T0 and the end of T3.

2.3 Mouse feeding study

For the mouse study, 16 pathogen-free male *FVB/N* mice were purchased from the NCI-Frederick Animal Production Area at 5 weeks of age. The mouse study was agreed to and regulated by the Animal Care and Use Committee of the National Cancer Institute - Frederick, MD (ASP-10-269). Mice were transferred to the Laboratory Animal Sciences Program of SAIC Frederick Inc. (Frederick, MD) and maintained throughout the study in single cages of 29.2×19.1×12.7 cm dimensions in a temperature (20–22 °C) and humidity (50%)-controlled room with a 12:12-hour light/dark cycle. Mice had *ad libitum* access to drinking water (reverse osmosis-purified water) and feed. Diets were prepared by Harlan Teklad (Madison, WI). All mice started on an AIN-93G purified diet [16], which was also the control diet. At 8.5 weeks of age, half of the mice were switched to a navy bean extract enriched diet. The control and the bean extract diet were the same except that part of the cornstarch (10% of the diet as fed) was replaced by navy bean ethanol extract, the preparation of which is described in Supporting Information S1.

Samples of the navy bean extract, control diet, and the bean extract diet were sent for metabolomic analysis. After being on the diet for 6 weeks, fecal samples as well as blood samples for serum analysis were taken from each mouse. We randomly selected 6 mice per treatment group for further fecal and serum analysis.

2.4 Serum assays

All serum, fecal, and diet samples were stored at <-70 °C until analysis. Metabolite profiles were measured at Metabolon Inc. (Research Triangle Park, North Carolina) using 3 different systems (liquid chromatography/mass spectroscopy +/- electrospray ionization and gas chromatography/mass spectroscopy). Separated metabolites were identified based on retention time, *m/z*, and the MS/MS spectral data of experimental data and >2,000 commercially available, purified, and authenticated standard compounds that were run across Metabolon's systems. Samples were run in a single run in batches of 30 samples. Blinded quality control samples of pooled serum were inserted at a level of 10% randomly throughout each batch (only for the PPT study), in addition to an unblinded standard every sixth sample (for all studies). The inter-assay and intra-assay coefficient of variation (CV) for metabolites of the blinded quality control sample were 9.7% and 15.6%, respectively. The inter-assay CV for internal metabolite standards was 5% in the PPT study and 6% in the LIFE and mouse study. The inter-assay CV for endogenous metabolites was 10% in the

LIFE study, 16% in the mouse study, and 12% in the PPT study, which are typical CVs for Metabolon analyses.

2.5 Statistical analyses

Parametric and nonparametric tests were used in Statistical Analysis Systems, version 9.2 (SAS, Inc.), software and STATA 9 (StatCorp, College Station, TX). The statistical analysis for each data set was restricted to metabolites that could be detected in 90% of samples after dry bean exposure in that specific data set. Signal intensity areas below the LOD were substituted with the minimum observed value in that specific data set data multiplied by 0.9. Peak intensities were not transformed.

To identify biomarkers of dry bean intake, serum metabolites before and after 4 weeks of high dry bean consumption were measured using samples from participants from the controlled human feeding study. Data were analyzed as fold-change (post-high dry bean diet as a percentage of pre-high dry bean diet) using paired *t*-tests and Wilcoxon signed-rank tests. We also determined the percentage overlap between pre-bean and post-bean samples of the study population. For the sensitivity analysis, we examined whether recent history of colorectal adenoma, insulin resistance, or feeding period modified the effects of dry bean consumption.

To confirm the identified biomarkers were specific to dry beans, we measured serum and fecal metabolites of mice that were on diets for 6 weeks that differed only in dry bean extract content. We statistically compared the 0% and 10% dry bean extract groups using unpaired *t*-tests (assuming unequal variance) and Mann-Whitney U-tests. The small sample size (*n*=6) and the non-normal distribution limited the statistical power; thus, the primary criterion for a specific serum biomarker or a specific fecal biomarker was that the peak areas of the mouse samples did not overlap between the two diets.

To determine if the identified biomarkers reflect dry bean consumption over several years, we conducted a nested study in the PPT. Serum metabolites and self-reported dry bean intake were measured at T0 and T3 for participants that reported increased dry bean consumption (*n*=106) and matched participants with unchanged, low dry bean consumption (*n*=106). To evaluate whether the identified biomarkers can detect long-term changes in dry bean intake within participants, we analyzed fold changes in serum metabolite levels (T3 as % of T0) in participants that increased their dry bean consumption. To evaluate whether the identified biomarkers can differentiate between participants with low and high self-reported dry bean intake, we analyzed differences in serum metabolite levels (high reporter T3 as % of low reporter) in matched participants with high or low dry bean consumption.

To determine whether changes in serum metabolite levels (T3 as % of T0) are gradual over a wide range of self-reported changes in dry bean intakes (T3-T0), we computed Pearson and Spearman correlation coefficients, the Kruskal Wallis test and multiple linear regression models using data from the nested study and included 21 participants, which did not fit into either category. The same statistical tests and data were used to determine whether the identified biomarker (T3) reflect differences in dry bean consumption over a wide range of intakes (T3). The median levels of each dry bean intake quantile and serum metabolite

quantile were used to determine correlation coefficients between intake and metabolite quantiles. For all multiple regression models, the following potential confounders were added to the models in a stepwise fashion and tested for their effect on serum metabolites: age, body mass index (BMI), and sex.

To adjust for multiple comparisons, q-values were calculated using the linear step-up procedure of Benjamini [17]. All statistical tests were two-sided and considered significant at $q < 0.05$.

3 Results

3.1 Controlled human feeding study (LIFE) identifies pipecolic acid as best indicator of dry bean consumption

A total of 275 known serum metabolites were identified; among these, 228 were above the LOD in 90% of post bean samples. After a 4-week dry bean-enriched diet (250 g/d), levels of 80 serum metabolites (35% of analyzed metabolites) were altered at $P < 0.05$ and 64 metabolites (28%) were altered at $q < 0.05$ using paired t-tests. The strongest responses to dry beans ($q < 10^{-7}$) were observed for pipecolic acid (PA), S-methyl-cysteine (SMC), N-acetylmethionine (NAO), trigonelline (TRIG), and indole propionate (IPA; Table 1; Figure 1). Serum PA increased in all participants after dry bean consumption. Without dry bean consumption, serum SMC and NAO could not be consistently detected. Recent history of colorectal adenoma, insulin resistance, or feeding period did not modify the effects of dry bean consumption.

3.2 Mouse feeding study identifies pipecolic acid and S-methyl-cysteine as best indicators of dry bean consumption

The objective of the mouse feeding study was to confirm that the observed changes in serum metabolites observed after consuming a dry-bean enriched diet in the LIFE study were in fact caused by dry beans and not a food component that was also changed by the recipes for the dry-bean enriched diet. Serum, fecal, and dietary metabolites were measured.

A total of 299 known serum metabolites were identified; among these, 285 were above the LOD in at least 5 of 6 mice on the bean extract diet. A navy bean extract-enriched diet altered 26 metabolites (9% of analyzed metabolites) at $P < 0.05$ and 7 metabolites (2.4%) at $q < 0.05$ using unpaired t-tests; the most significant being PA, SMC, and NAO (all $q < 0.003$). Besides isovalerylcarnitine, PA, SMC and NAO separated mice from the control and the bean extract diet. All mice fed bean extract had significantly higher PA (median: 5,062% of control mice), SMC (217%) and NAO (2,627%) levels than control mice (Figure 2). Without dry bean consumption, serum SMC and NAO could not be consistently detected. Serum TRIG and IPA were below the LOD in all samples.

A total of 370 known fecal metabolites were identified; of these, 302 were above the LOD in at least 5 of 6 mice on the bean extract diet. A navy bean extract enriched diet altered 49 metabolites (16% of analyzed metabolites) at $P < 0.05$ and 10 metabolites (3.3%) at $q < 0.05$ using unpaired t-tests. Seventeen fecal metabolites, including PA and SMC, separated mice from the control and the bean extract diet. All mice fed bean extract had higher PA (median:

704% of control mice) and SMC (139%) levels than control mice (Figure 2). Without dry bean consumption, serum levels of SMC could not be consistently detected. All fecal NAO and IPA levels were below the LOD and TRIG values could not differentiate the groups (data not shown).

Based on the peak area, the most abundant metabolites in the dry bean extract were PA and TRIG, which constituted 37.1% and 10.1% of the combined peak area of all known metabolites. In the BE supplemented diet, they made up 35.3% (PA) and 8.2% (TRIG) of the combined peak area. In comparison, the PA and TRIG peaks constituted 0.12% and 0.03% of the combined peak area in the control diet. The most abundant peaks unique to the BE supplemented diet were the SMC (1.37%) and NAO peaks (0.66%). In the dry bean extract, they made up 0.75% (SMC) and 0.32% (NAO) of the combined peak area. IPA was not detected in dietary samples.

3.3 Self-Reported Dry Bean Consumption (Polyp Prevention Trial) identifies pipecolic acid and S-methyl-cysteine as best indicators of dry bean consumption

The objective was to determine whether the identified metabolites reflect dry bean consumption over several years. We used a nested study of polyp-free participants within the PPT, a 4-year intervention study that promoted dry bean consumption in the intervention arm.

A total of 454 known serum metabolites were identified; of these, 305 were above the LOD in 90% of T0 or T3 samples. Increased self-reported dry bean consumption from T0 by at least 31 g/d to 42 g/d in year 3 was associated with changes in 77 serum metabolites (25% of analyzed metabolites) at $P < 0.05$ and 42 metabolites (14%) at $q < 0.05$ using paired t-tests. The 42 metabolites included PA, SMS, NAO, and TRIG (Table 2; Figure 3). Serum levels of PA, SMC, NAO, TRIG, and IPA did not change at $P > 0.10$ in participants with low, unchanged self-reported dry bean consumption (data not shown).

Self-reported high or low dry bean consumption in year 3 was associated with differences in 37 serum metabolites (12% of analyzed metabolites) at $P < 0.05$ and 4 metabolites (1.3%) at $q < 0.05$ using a paired t-test. Besides dihomolinolenate, the 4 metabolites included PA, SMC, and IPA (Table 3; Figure 3).

Changes in self-reported dry bean intake over a 3 year period were linearly reflected in changes in serum PA, SMC, NAO, TRIG, and IPA levels during the same time period (Table 2; Figure 4). Using six groups that span changes in self-reported dry bean intake (T3-T0; <0, 0, 0.5–14, 20–43, 49–87, and 93–289 g/d) was highly correlated with changes in serum PA, SMC, NAO, TRIG, and IPA levels (all $r > 0.85$).

Differences in self-reported dry bean intake in year 3 were linearly reflected in serum PA ($r=0.34$) and SMC ($r=0.43$) levels, to a smaller extent in NAO and IPA levels, but not in TRIG levels (Table 3; Figure 4). TRIG levels, however, were associated with coffee consumption ($r=0.55$; results not shown). Using six groups that span the self-reported dry bean intake in year 3 (0, 1–5, 7–28, 42–72, 98–108, and 147–294 g/d), dry bean consumption was correlated strongly with serum PA ($r=0.97$), SMC ($r=0.96$), and IPA

($r=0.87$) levels, less correlated with NAO levels ($r=0.47$), and not correlated with TRIG levels ($r=0.11$). Age, sex, and BMI did not affect the estimates for any of the five examined metabolites, except for IPA, which was affected by BMI (lower in obese compared with overweight and normal BMI participants; data not shown).

4 Discussion

Accumulating evidence suggests that there is a protective role of dry bean intake against colorectal neoplasia occurrence and progression; however, most of the evidence in human studies relies on self-reported dietary assessment. To demonstrate a protective effect of dry beans on CRC in humans, biomarkers of dry bean intake are needed. Currently, there are few established dietary biomarkers of diet, and none for dry beans. Metabolomics provides a global analysis of metabolites in biological samples and, thus, is a well-suited analytical tool for identifying markers of dry bean consumption. Criteria for a good serum biomarker of dietary intake are: a) strong association with dietary intake or changes in dietary intakes that are consistent across various populations, b) abundance and relative specificity to dietary compound of interest, c) low intestinal microbial metabolism of biomarker and its precursor, d) high intestinal absorption rate, e) minimal endogenous synthesis or catabolism, and e) long half-life in serum. Dietary biomarkers may not necessarily be bioactive markers for prevention of colorectal neoplasia.

Identification and evaluation of markers of dry bean consumption was a three-phase process. We used three studies that each had shown chemopreventive effects of dry bean consumption on at least biomarkers of CRC risk [7–9, 18]. In a controlled human feeding study (LIFE) and a controlled mouse feeding study, we identified serum PA, SMC, NAO, and TRIG as markers of dry bean consumption; among those, SMC and NAO were specific to dry bean intake. In a multi-year intervention study that promoted dry bean consumption in the intervention arm (PPT) only serum PA and SMC separated individuals based on dry bean consumption.

4.1 Pipecolic acid

PA is a cyclic, non-protein imino acid, which is the most abundant non-protein nitrogen fraction in dry beans (6,788 mg/kg dry beans) [19–20]. In comparison, most commonly consumed foods have negligible PA contents (<2.6 mg/kg) except for cabbage (21.37 mg/kg), broccoli (13.56 mg/kg), and cauliflower (11.90 mg/kg) of the *Brassicaceae* family [20, 22].

Intestinal microbial metabolism, which converts lysine to PA, can affect concentrations of PA in biological samples and may explain the variability in magnitude of serum and fecal PA response to dry beans in our study. Prior studies reported that antibiotic treatment decreased plasma PA concentrations [21, 23]. Intestinal microbial metabolism of dietary PA may confer chemopreventive benefits for the human host because PA is a precursor of microbial compounds with anti-inflammatory (e.g., rapamycin), antitumor (e.g., swainsonine), and antibiotic (e.g., virginiamycin) properties [19]. Some rare, metabolic disorders can increase PA concentrations in blood [21, 23–26]. Taken together these findings indicate that serum PA is specific to dry bean intake.

4.2 S-Methyl cysteine

SMC is a non-protein amino acid, which is derived by acid hydrolysis from the second most common non protein nitrogen component of dry beans γ -glutamyl-SMC (3,026 mg/kg dry beans) [20]. Besides dry beans, only some other *Phaseolus* and *Vigna* species (mung beans, cow peas) accumulate sulfur groups as SMC components [27–28]. *Alliaceae* and *Cruciferae* vegetables contain small amounts of SMC [29–31]. The chemopreventive properties of dry beans may be partly conferred through SMC, as it inhibits inflammation [32], lipid peroxidation [33], and chemical carcinogenesis in the liver and colon [34–36] by acting as a methyl-group acceptor.

SMC has a high intestinal absorption rate [37–39] and negligible intestinal microbial SMC catabolism [40]. Endogenous catabolism of SMC is extensive (88%) but relatively slow as 41% is excreted within 24 hr as urine [37]. Under rare circumstances, SMC-containing proteins can be synthesized endogenously during xenobiotic metabolism [41, 42]. Taken together these findings indicate that serum SMC is specific to dry bean intake.

4.3 Strengths and limitations

There were several strengths and limitations of our studies for identifying and monitoring markers of dry bean intake. The greatest strength is that we used three nutrition studies that complement each other in their strengths and weaknesses. The LIFE study was a controlled human feeding study; however, changes in dry bean consumption were large (250 g of cooked dry beans/d), only men were enrolled, and changes in other food components associated with the high dry bean diet could have caused metabolite changes. The mouse feeding study was a controlled feeding study, in which diets differed only in dry bean content; however, only male mice were in the study. The nested study within the PPT was a multi-year intervention study that included polyp-free men and women over a wide intake range (0 to 300 g/d) of dry bean intake; however, intakes were based on self-reported data using a FFQ. Since only 3 participants in the control arm reported increasing their dry bean consumption by 31 g/d and 25 participants in the intervention arm had unchanged low dry bean consumption, we could not evaluate, whether the study arm affected the results for the PPT. Since only 5 participants in the high dry bean group of the PPT were smoker and smoking was an exclusion criterion for the LIFE, we could not evaluate the effect of smoking on serum metabolite levels.

A general limitation was that metabolomics is a global analysis tool and, therefore, cannot be as accurate as an optimized assay for determining specific metabolite levels; although the metabolite CVs from our three studies indicated good reliability.

Given that the PPT was based on self-reported, typical dry bean consumption for the previous year and the LIFE and mouse feeding studies were controlled feeding studies, only marginal associations between self-reported dry bean intake and serum markers that were identified in the LIFE and mouse feeding studies can be expected in the PPT. However, correlations of $r > 0.95$ between 6 groups that span self-reported dry bean intake with serum PA and SMC suggest that PA and SMC are useful biomarkers of dry bean consumption and that self-reported dry bean intake approximated “true” intake levels of dry beans. Similar

correlations as for SMC and dry bean intake by a FFQ for the previous year ($r=0.43$) have been reported for the linear association between biomarkers of α -tocopherol and 6 separate 24-hr dietary recalls [43]. The prospective collection of dietary data and the use of dietary questionnaires that were reviewed by registered dietitians may have improved the accuracy of the FFQ in this study [6].

4.4 Concluding remarks

We previously showed in a controlled human feeding study (LIFE), a large, multi-year nutrition intervention study (PPT), and a mouse feeding study that dry bean intake has chemopreventive effects for colorectal neoplasia. In the current study, we used the same three studies to identify and evaluate markers of dry bean consumption. We observed that PA and SMC were elevated after dry bean consumption in the LIFE study and a mouse study and reflected dry bean consumption in the Polyp Prevention Trial. Serum PA and SMC levels are specific to dry beans; therefore, PA and SMC can be used as metabolite markers for dry bean consumption. Future research will determine whether PA and SMC are bioactive marker for prevention of colorectal neoplasia.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding: The authors would like to thank the Polyp Prevention Trial Study Group for their outstanding contribution to this project. In addition, the authors thank Edward Karoly, Elizabeth Morin-Kensicki, Robert Mohny, Jeff Pfohl, and Brante Sampey from Metabolon (Research Triangle Park, North Carolina) for metabolomics analysis of the serum samples; Jen Wise and Craig Driver of the Laboratory Animal Sciences Program of SAIC Frederick Inc. for taking care of the mice; and Chris Dextras and Thomas G. McCloud from the Natural Products Support Group of Applied Developmental Research Program of SAIC Frederick Inc. for preparing the dry bean extract.

This study was funded by the Office of Complementary and Alternative Medicine, the Office of Dietary Supplements, the US Dry Bean Council, Oregon State University, the General Research Center at Penn State University (NIH grant M01 RR10732), and the Intramural Research Program, National Cancer Institute, NIH, DHHS, Bethesda, MD (Z01 CP10198-08; ZIA BC 010025; subcontract 25XS101).

Abbreviations

BMI	Body Mass Index
CRC	Colorectal Cancer
FFQ	Food Frequency Questionnaire
IPA	Indolepropionate
IQR	Interquartile Range
LIFE	Legume Inflammation Feeding Experiment
NAO	N-Acetylmethionine
PA	Pipecolic Acid

PPT	Polyp Prevention Trial
SMC	S-Methyl-Cysteine
TRIG	Trigonelline

References

1. Ferlay J, Soerjomataram I, Ervik M, Dikshit R, et al. Cancer incidence and mortality worldwide: IARC Cancer base No. 11. GLOBOCON. 2013
2. U.S. Cancer Statistics Working Group. United States cancer statistics: 1999–2011 incidence and mortality web-based report. Department of Health and Human Services, Centers for Disease Control and Prevention, and National Cancer Institute; 2014.
3. Mariotto AB, Yabroff KR, Shao Y, Feuer EJ, et al. Projections of the cost of cancer care in the United States: 2010–2020. J. Natl. Cancer. Inst. 2011; 103:117–128. [PubMed: 21228314]
4. World Cancer Research Fund/American Institute for Cancer Research. Food nutrition physical activity and the prevention of cancer: A global perspective. Washington DC: AICR; 2007.
5. Wang Y, Wang Z, Fu L, Chen Y, et al. Legume consumption and colorectal adenoma risk: a meta-analysis of observational studies. PloS One. 2013; 8:e67335. [PubMed: 23826270]
6. Lanza E, Hartman TJ, Albert PS, Shields R, et al. High dry bean intake and reduced risk of advanced colorectal adenoma recurrence among participants in the polyp prevention trial. J. Nutr. 2006; 136:1896–1903. [PubMed: 16772456]
7. Bobe G, Murphy M, Rogers CJ, Hance KW, et al. Serum adiponectin, leptin, C-peptide, homocysteine, and colorectal adenoma recurrence in the Polyp Prevention Trial. Cancer Epidemiol. Biomarkers Prev. 2010; 19:1441–1452. [PubMed: 20501764]
8. Bobe G, Murphy G, Albert PS, Sansbury LB, et al. Serum cytokine concentrations, flavonol intake and colorectal adenoma recurrence in the Polyp Prevention Trial. Br. J. Cancer. 2010; 103:1453–1461. [PubMed: 20924374]
9. Hartman TJ, Albert PS, Zhang Z, Bagshaw D, et al. 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. 2010; 140:60–67. [PubMed: 19889807]
10. Zhang Z, Lanza E, Kris-Etherton PM, Colburn NH, et al. A high legume low glycemic index diet improves serum lipid profiles in men. Lipids. 2010; 45:765–775. [PubMed: 20734238]
11. Bobe G, Sansbury LB, Albert PS, Cross AJ, et al. Dietary flavonoids and colorectal adenoma recurrence in the Polyp Prevention Trial. Cancer Epidemiol. Biomarkers Prev. 2008; 17:1344–1353. [PubMed: 18559549]
12. O’Gorman A, Gibbons H, Brennan L. Metabolomics in the identification of biomarkers of dietary intake. Comput. Struct. Biotechnol. J. 2013; 4:1–7.
13. Hedrick VE, Dietrich AM, Estabrooks PA, Savla J, et al. Dietary biomarkers: advances, limitations and future directions. Nutr. J. 2012; 11:2–14. [PubMed: 22230384]
14. Schatzkin A, Lanza E, Corle D, Lance P, et al. Lack of effect of a low-fat, high-fiber diet on the recurrence of colorectal adenomas. Polyp Prevention Trial study group. N. Engl. J. Med. 2000; 342:1149–1155. [PubMed: 10770979]
15. Lanza E, Schatzkin A, Daston C, Corle D, et al. Implementation of a 4-y, high-fiber, high-fruit-and-vegetable, low-fat dietary intervention: results of dietary changes in the Polyp Prevention Trial. Am. J. Clin. Nutr. 2001; 74:387–401. [PubMed: 11522565]
16. Reeves P, Nielsen F, Fahey GJ. AIN-93 purified diets for rodents: final report of the American Institute of Nutrition ad hoc writing committee on the reformulation of the AIN-76A rodent diet. J. Nutr. 1993; 123:1939–1951. [PubMed: 8229312]
17. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A practical and powerful approach to multiple testing. J. R. Stat. Soc. B. 1995; 57:289–300.

18. Lanza E, Schatzkin A, Ballard-Barbash R, Corle D, et al. The polyp prevention trial II: dietary intervention program and participant baseline dietary characteristics. *Cancer Epidemiol. Biomarkers Prev.* 1996; 5:385–392. [PubMed: 9162305]
19. He M. Pipecolic acid in microbes: biosynthetic routes and enzymes. *J. Ind. Microbiol. Biotechnol.* 2006; 33:401–407. [PubMed: 16418868]
20. Zacharius RM. Composition of the nitrogenous components of the bush bean seed (*Phaseolus vulgaris*) including isolation of δ -N-acetylmethionine. *Phytochemistry.* 1970; 9:2047–2051.
21. Fujita T, Hada T, Higashino K. Origin of D- and L-pipecolic acid in human physiological fluids: a study of the catabolic mechanism to pipecolic acid using the lysine loading test. *Clin. Chim. Acta.* 1999; 287:145–156. [PubMed: 10509903]
22. Fujita T, Fujita M, Kodama T, Hada T, et al. Determination of D- and L-pipecolic acid in food samples including processed foods. *Ann. Nutr. Metab.* 2003; 47:165–169. [PubMed: 12743469]
23. Matsuda Y, Fujita T, Hada T, Higashino K. Comparative study on the correlation of plasma γ -aminobutyric acid and pipecolic acid with liver function in patients with liver cirrhosis. *Hepatol. Res.* 2000; 18:132–140. [PubMed: 10936564]
24. Chang YF. Lysine metabolism in the human and the monkey: demonstration of pipecolic acid formation in the brain and other organs. *Neurochem. Res.* 1982; 7:577–588. [PubMed: 6811962]
25. Broquist HP. Lysine-pipecolic acid metabolic relationships in microbes and mammals. *Ann. Rev. Nutr.* 1991; 11:435–448. [PubMed: 1909881]
26. Peduto A, Baumgartner MR, Verhoeven NM, Rabier D, et al. Hyperpipecolic acidemia: a diagnostic tool for peroxisomal disorders. *Mol. Genet. Metab.* 2004; 82:224–230. [PubMed: 15234336]
27. Evans IM, Boulter D. S-Methyl-L-cysteine content of various legume meals. *Qual. Plant. Plant Food Hum. Nutr.* 1975; 24:257–261.
28. Bell EA, Watson AA, Nash RJ. Non-protein amino acids: a review of the biosynthesis and taxonomic significance. *Nat. Prod. Commun.* 2008; 3:93–110.
29. Virtanen AI. Some organic sulfur compounds in vegetables and fodder plants and their significance in human nutrition. *Angew. Chem. Int. Ed. Engl.* 1962; 1:299–306.
30. Yamazaki Y, Iwasaki K, Mikami M, Yagihashi A. Distribution of eleven flavor precursors, S-alk(en)yl-L-cysteine derivatives, in seven *Allium* vegetables. *Food Sci. Technol. Res.* 2010; 17:55–62.
31. Kubec R, Drhová V, Velíšek J. Thermal degradation of S-methylcysteine and its sulfoxide important flavor precursors of Brassica and Allium vegetables. *J. Agr. Food Chem.* 1998; 46:4334–4340.
32. Yin MC, Hsu CC, Chiang PF, Wu WJ. Antiinflammatory and antifibrogenic effects of s-ethyl cysteine and s-methyl cysteine in the kidney of diabetic mice. *Mol. Nutr. Food Res.* 2007; 51:572–579. [PubMed: 17440992]
33. Higuchi O, Tateshita K, Nishimura. Antioxidative activity of sulfur-containing compounds in Allium species for human low-density lipoprotein (LDL) oxidation in vitro. *J. Agr. Food Chem.* 2003; 51:7208–7214. [PubMed: 14611195]
34. Wei M, Wanibuchi H, Yamamoto S, Iwai S, et al. Chemopreventive effects of S-methylcysteine on rat hepatocarcinogenesis induced by concurrent administration of sodium nitrite and morpholine. *Cancer Lett.* 2000; 161:97–103. [PubMed: 11078918]
35. Fukushima S, Takada N, Hori T, Min W, et al. Suppression of chemical carcinogenesis by water-soluble organosulfur compounds. *J. Nutr.* 2001; 131:1049S–1053S. [PubMed: 11238814]
36. Nishikawa-Ogawa M, Wanibuchi H, Morimura K, Kinoshita A, et al. N-acetylcysteine and S-methylcysteine inhibit MeIQx rat hepatocarcinogenesis in the post-initiation stage. *Carcinogenesis.* 2006; 27:982–988. [PubMed: 16338951]
37. Mitchell SC, Smith RL, Waring RH, Aldington GF. The metabolism of S-methyl-L-cysteine in man. *Xenobiotica.* 1984; 14:767–779. [PubMed: 6506751]
38. Waring RH, Harris RM, Steventon GB, Mitchell SC. Degradation to sulphate of S-methyl-L-cysteine sulfoxide and S-carboxymethyl-L-cysteine sulfoxide in man. *Drug Metab. Drug Interact.* 2003; 19:241–256.

39. Edmands WM, Gooderham NJ, Holmes E, Mitchell SC. S-Methyl-L-cysteine sulphoxide: the Cinderella phytochemical? *J. Toxicol. Res.* 2013; 2:11–22.
40. Schwiertz A, Deubel S, Birringer M. Bioactivation of selenocysteine derivatives by β -lyases present in common gastrointestinal bacterial species. *Int. J. Vitam. Nutr. Res.* 2008; 78:169–174. [PubMed: 19326339]
41. Garnier R, Rambourg-Schepens MO, Müller A, Hallier E. Glutathione transferase activity and formation of macromolecular adducts in two cases of acute methyl bromide poisoning. *Int. J. Occup. Environ. Med.* 1996; 53:211–215.
42. Rubino FM, Pitton M, Di Fabio D, Meroni G. Measurement of S-methylcysteine and S-methyl-mercapturic acid in human urine by alkyl-chloroformate extractive derivatization and isotope-dilution gas chromatography–mass spectrometry. *Biomed. Chrom.* 2011; 25:330–343.
43. Lebold KM, Ang A, Traber MG, Arab L. Urinary α -carboxyethyl hydroxychroman can be used as a predictor of α -tocopherol adequacy, as demonstrated in the Energetics Study. *Am. J. Clin. Nutr.* 2012; 96:801–809. [PubMed: 22952171]

Figure 1.

A 4-week dry bean-enriched diet (250 g/d) increased abundance of serum pipecolic acid, S-methyl-cysteine, N-acetylornithine, and trigonelline in 46 middle-aged men in a controlled human dry bean feeding study (LIFE). On the left side, crosses represent values of individual participants and horizontal, solid lines refer to median and interquartile ranges. On the right side, box and whiskers plot are shown; the horizontal, solid lines refer to median and interquartiles, the plus signs represent the mean, dots represent values of individual participants that were beyond 1.5 times the interquartile distance, and the

whiskers go 1.5 times the interquartile distance or to the highest or lowest point, whichever is shorter.

6-Week Mouse Feeding Study

Figure 2.

A 6-week 10% dry bean extract-containing diet increased abundance of serum and fecal pipercolic acid and S-methyl-cysteine, as well as serum N-acetylornithine in a mouse feeding study (n = 6 mice per group). Dots and crosses represent values of individual control and dry bean extract fed mice, respectively. Horizontal, solid lines refer to median and interquartile ranges.

3-Year Human Nutrition Study – Self Report for Previous Year (PPT)

Figure 3.

Association between self-reported dry bean consumption in year 3 and serum pipecolic acid, S-methyl-cysteine, N-acetylornithine, and trigonelline at the end of year 3 among polyp-free men and women in the Polyp Prevention Trial (PPT). High dry bean consumers (n = 106) increased their self-reported dry bean intake from year 0 by at least 31 g/d to 42 g/d in year 3, whereas low dry bean consumers (n = 106) self-reported an unchanged dry bean consumption of <8 g/d. The description of symbols and lines is provided in Figure 1.

Figure 4.

Linear association between self-reported dry bean consumption in year 3 and serum pipecolic acid, S-methyl-cysteine, N-acetylornithine, and trigonelline at the end of year 3 among polyp-free men and women in the Polyp Prevention Trial (PPT). Dots represent values of individual participants and the solid lines represent the regression lines.

Table 1Serum metabolites most affected by dry bean consumption (LIFE study)^a

Metabolite	Change (Post as % of Pre)		Change (% of Participants)	
	Median	IQR ^b	Increased	Decreased
Pipecolic acid	+660	+421, +1,190	100	0
S-Methyl-cysteine	+233	+103, +610	89	4
N-Acetylmethionine	+85	+44, +149	91	4
Trigonelline	+140	+77, +280	93	7
Indole propionate	+99	+54, +99	89	9

^a) Overnight fasted blood samples were collected from 46 participants at the beginning and end of a 4-week high dry bean diet (250 g/d). Only metabolites are shown with $q < 10^{-7}$ (<10% of participants changed in opposite direction to majority).

^b) IQR: interquartile range.

Table 2Effect of self-reported changes in dry bean consumption on selected metabolites (PPT study)^a

Metabolite	Change (T3 as % of T0)		Correlation with Dry Bean Intake Change (T3-T0)	
	Median	IQR ^b	q-Value	Quantiles
Pipecolic acid	+44	-11, +135	0.0005	0.30
S-Methyl-cysteine	+27	-16, +67	0.01	0.40
N-Acetylornithine	+15	-8, +41	0.006	0.24
Trigonelline	+22	-24, +143	0.02	0.22
Indole propionate	+32	-14, +120	0.06	0.17
				0.90

^a) Overnight fasted blood samples were collected at baseline (T0) and at the end of year 3 (T3) from 106 participants who increased their dry bean intake (+31 g/d from T0) to 42 g/d in year 3. Only the 5 metabolites identified in the controlled human feeding study (LIFE study) are shown. For continuous, we correlated changes in self-reported dry bean intake with changes in serum metabolite levels. For quantiles, we correlated median changes of six quantiles of self-reported dry bean intake (T3-T0; <0, 0, 0.5–14, 20–43, 49–87, and 93–289 g/d) with median changes in metabolite levels of the quantiles.

^b) IQR: interquartile range.

Table 3
Effect of self-reported dry bean consumption on selected metabolites (PPT study)^a

Metabolite	Difference (High as % of Low at T3)		Correlation with Dry Bean Consumption (T3)	
	Median	IQR ^b	q-Value	Continuous Quantiles
Pipecolic acid	+61	-14, +171	0.003	0.34 0.97
S-Methyl-cysteine	+36	-16, +111	0.0001	0.43 0.96
N-Acetylmethionine	+15	-30, +80	0.24	0.19 0.47
Trigonelline	+14	-40, +221	0.93	0.0006 0.11
Indole propionate	+73	-21, +207	0.02	0.21 0.87

^a Overnight fasted blood samples, collected at baseline (T0) and at the end of year 3 (T3), were compared from Participants (n=106 per group) self-reported either increased (+31 g/d T3-T0), high dry bean intake of 42 g/d (T3) or low, unchanged dry bean consumption of <8 g/d at T0 and T3. Only the 5 metabolites identified in the controlled human feeding study (LIFE study) are shown. For continuous, we correlated changes in self-reported dry bean intake with changes in serum metabolite levels. For quantiles, we correlated median values of six quantiles of self-reported dry bean intake (T3; 0, 1–5, 7–28, 42–72, 98–108, and 147–294 g/d) with median metabolite levels of the quantiles.

^b IQR: interquartile range