Search for novel circulating cancer chemopreventive biomarkers of dietary rice bran intervention in *Apc*^{Min} mice model of colorectal carcinogenesis, using proteomic and metabolic profiling strategies.

Authors

Leonie Norris[1+], Aditya Malkar[2+], Emma Horner-Glister[1], Amirmansoor Hakimi[1], Leong L Ng[3], Andreas J Gescher[1], Colin Creaser[1], Stewart Sale[1*] and Donald JL Jones[1,3*].

Addresses:

[1]Department of Cancer Studies and Molecular Medicine, Leicester Royal Infirmary, University of Leicester, Leicester. LE2 7LX.

[2]Centre for Analytical Science, Department of Chemistry, Loughborough University, Loughborough, Leicestershire LE11 3TU, UK.

[3]Department of Cardiovascular Sciences and NIHR Leicester Cardiovascular Biomedical Research Unit, Glenfield Hospital, Groby Road, Leicester, LE3 9QP, UK.

*Corresponding author Dr Donald JL Jones djlj1@le.ac.uk

[+] These authors contributed equally to the work

Abstract

Scope:

There is strong epidemiological evidence indicating that consumption by humans of whole-grain foods including rice bran may be associated with a low incidence of cancer, especially in the colorectum. Molecular processes associated with cancer development may be retarded by fibre consumption. Consequently intervention with dietary fibre might be suitable as a cancer chemoprevention strategy in high-risk populations. Here we searched for putative molecular mechanism-based efficacy biomarkers of rice fibre consumption in the plasma of mice characterised by a genetic propensity to develop gastrointestinal adenomas. The hypothesis was tested that metabolic and proteomic changes in blood reflect the chemopreventive activity of rice bran.

Methods and Results:

ApcMin mice received diet supplemented with rice bran at 5, 15 and 30%. Blood and tissue samples were taken. Plasma was subjected to mass spectrometry-based proteomic and metabolic profiling analyses as well as assessment of haematocrit values. Gastrointestinal tracts were removed and adenomas were counted and their size was measured so that total tumour burden could be calculated. The hypothesis was tested that metabolic and proteomic changes in blood reflect chemopreventive activity.

Conclusion:

Rice bran consumption reduced adenoma burden and number in a dose-related fashion when compared to controls. Metabolic profiling data demonstrated strong clustering of the groups indicating that metabolic pathways are perturbed. Proteomic analysis identified adiponectin as a molecule that was significantly altered, which may play a role in tumour suppression.

1.0 Introduction

Epidemiological evidence suggests that human consumption of whole-grain foods is associated with a low incidence of cancer, especially in the colorectum [1]. Rice is the staple food of over half the world's population. Whilst the unpolished brown (bran-containing) variety possesses special dietary importance in Asia, rice consumed in the Western world is generally white, obtained from brown rice by bran removal. Dietary differences such as this may explain why the incidence of certain cancers, including those in the colorectum, is much lower in Asia than in the West [2]. A stabilised rice bran preparation containing 25-35% fibre, when added at 30% to the diet, interfered with adenoma development in the Apc^{Min} min mouse, a model of colorectal carcinogenesis [3]. These mice possess a heritable *Apc* gene mutation which is responsible for adenoma formation [4]. Importantly, the fibre content was pivotal for the adenoma retarding activity of the rice bran, as a rice bran formulation from which most of the fibre had been removed was devoid of efficacy in this model [3]. Three types of mechanism have been proffered by which dietary fibre including that from rice is thought to potentially interfere with colorectal carcinogenesis. Firstly, fibre is thought to exert 'physical' effects such as increasing faecal bulk, hastening faecal transit and binding potentially co-carcinogenic bile salts. Secondly, fibre may beneficially change the gastrointestinal microflora and luminal environment. Thirdly, fibre carbohydrates can undergo fermentation in the luminal environment, generating short chain fatty acids such as butyrate, which are thought to exert anticarcinogenesis [3]. Whilst the evidence for the anticarcinogenic properties of fibre particularly against colorectal cancer is tantalising, details of the underlying molecular processes remain obscure. A major shortcoming in the clinical optimisation of cancer chemopreventive interventions is the scarcity of knowledge of circulating markers which signal chemopreventive efficacy early during the intervention period. Proteomics and metabolomics are two attractive methodological platforms which help identify such biomarkers [5] revealing specific pathways pertinent to carcinogenic and/or

anticarcinogenic mechanisms [6-9]. Notably, a ^{1}H-NMR metabolomic approach recently identified an increase in the phosphocholine/glycerophosphocholine ratio, a biochemical marker of *Apc* tumour suppressor-mediated immortalisation and malignant transformation of cells, in the small bowel and colon of Apc^{Min} mice [10]. We wished to identify biochemical perturbations which correlate with rice bran fibre-mediated anti-colorectal carcinogenesis. To that end, the plasma of Apc^{Min} mice which received rice bran with their diet was subjected to proteomic and metabolic profiling analyses. The ultimate aim of the work was to buttress the preclinical knowledge of the potential usefulness of rice bran as a cancer chemopreventive intervention.

2.0 Methods and materials

2.1 Reagents

HPLC grade water, HPLC grade acetonitrile, 98% molecular biology dithiothreitol, purified grade ethylenediaminetetraacetic acid (EDTA) and reagent grade iodoacetamide were purchased from Sigma-Aldrich (Poole, Dorset, UK). Ultra grade (≥99.5%) ammonium bicarbonate was purchased from Fluka (Buchs, Switzerland). Mass spectrometry grade Trypsin Gold was purchased from Promega (Madison WI). Enolase digestion standard (S. cerevisiae) and alcohol dehydrogenase digestion standard (S. cerevisiae) were purchased from Waters Corp. (Milford, MA). Lo-Bind sample tubes were purchased from Eppendorf (Hamburg, Germany). LCMS grade water, acetonitrile (for metabolic profiling) and formic acid was purchased from Fisher Scientific (Loughborough, UK). Trasylol was purchased from Bayer (Newbury, UK). The rice bran formulation used ("RiFiber") was generously provided by NutraCea (now RiceBran Technologies, Scottsdale, Arizona, USA). The fibre content of RiFiber is between 40-50% of total weight, of this fibre 1% or less is soluble, so that the ratio of soluble to insoluble fibre in the fibre fraction is ~1:100.

2.2 Animal Study:

Apc^{Min} mice received standard powdered AIN93G [11] rodent diet supplemented with RiFiber at 5, 15 or 30 % (n=15 per treatment group) and the Apc^{Min} genotype was confirmed by PCR. The daily intake of rice fibre by each mouse was approximately 63, 189 or 378 mg, respectively per diem. Experiments were performed under animal project licence PPL 80/2167, granted to the University of Leicester by the UK Home Office. The experimental design was assessed by the University of Leicester Local Ethical Committee for Animal Experimentation and met the standards set by the UKCCCR guidelines. Mice received this diet from 4 weeks of age until termination of study at 16 weeks of age. At the termination of study, mice were culled by cardiac exsanguination and blood collected for proteomic analysis. No mice died prematurely during the study. Blood was collected into Li-Heparin tubes (Sarstedt, Leicester). Tubes were centrifuged at 3000rpm for 10mins and the plasma portion removed, aliquoted and stored at -80°C. Gastrointestinal tracts were removed and washed with PBS to remove all food content and then were opened longitudinally to reveal adenomas. Adenomas were counted and their sizes were measured using digital callipers so that total tumour burden could be calculated. An aliquot of blood from each animal was used to measure haematocrit value.

2.3 Proteomics

2.3.1 Instrumentation

Discovery

Samples were analysed on a Waters NanoAcquity system (Waters Corporation, Milford, MA). The peptides were initially loaded onto a Symmetry C_{18} 180 µm x 20 mm, 5 µm trap column to desalt and chromatographically focus the peptides prior to elution onto a HSS T3 C_{18} 75 µm x 150 mm, 1.7 µm analytical column (Waters Corporation, Milford, MA). Solvent A: HPLC grade water with 0.1% formic

acid and Solvent B: Acetonitrile with 0.1% formic acid was used. The flow rate was set at 0.3 µL/min. The gradient began following a 3 min (5 µl/min) trapping stage on the trap column. At time zero, B was 1%. B increased linearly to 40% at 90min and to 85% at 92 min. The gradient was held at 85% B at 93 min and returned to starting conditions at 95 min to equilibrate. Total run time was 110min.

The LC system was coupled to a Waters Synapt G2 HDMS (Waters Corporation, Milford, MA). The instrument was operated in positive ion nanoelectrospray ionisation mode. Capillary voltage was set at 2.8Kv and cone voltage at 30V. Picotip emitters (10 µm internal diameter, Presearch, Basingstoke, UK) were used for the nanostage to direct flow from the analytical column through to the source. A helium gas flow of 180mL/min and ion mobility separator gas flow (N_2) of 90 mL/min with a pressure of 2.5 mbar was used. An IM wave velocity of 600 m/s and wave height of 40V was used throughout each run. During LC-IM-Data Independent acquisition (DIA)-MS low CID energy, 2V was applied across the transfer ion guide. During high CID energy, a ramp of 27-50V was applied. Argon was used as CID gas. Lockspray of [Glu1]-Fibrinopeptide (GFP) *m/z* at 785.84265 was used to maintain mass accuracy throughout the chromatographic run. Data were acquired in $HDMS^E$ mode from m/z 100-1950 with a scan speed 0.9s/scan using MassLynx 4.1 (Waters Corporation (Milford, MA)).

Targeted analysis of candidate proteins

Targetted analysis of proteins was carried out using LC-MS/MS-Selected reaction monitoring (SRM) on a Waters Xevo TQ coupled to a nanoacquity nanoUPLC. The SRM transitions were extracted from the data collected on the G2. VerifyE enabled the creation of several SRM transitions using the exact mass –retention time (EMRT) tags for each peptide. A minimum of 3 peptides were used for each protein. The proteins were then quantified relatively by comparing the mean average of the MS1 intensity of the top 3 peptides [12,13]. Students t-test analysis was conducted for all samples using GraphPad Prism 6.0 to test the difference between control and 30%.

The same chromatography and source conditions used in the discovery phase on the Waters Synapt G2 were transferred to the Waters Xevo TQ, thus ensuring consistency in the transition between discovery and verification steps.

2.3.2 Sample Preparation

Immunodepletion

Discovery phase plasma samples (10 µl) from 3 pooled male and 3 pooled female mice from both control and 30 % rice bran treated were immunodepleted. Each pooled sample was made up from 3-4 individual mice. The top 20 high abundant proteins in plasma were depleted using ProteoPrep 20 immunodepletion spin columns (Sigma-Aldrich, Poole, UK) containing antibodies against: Albumin, Immunoglobulin G, A, M and D, Transferrin, Fibrinogen, α-2-Macroglobulin, α-1-Antitrypsin, Haptoglobin, α-1 Acid Glycoprotein, Ceruloplasmin, Apolipoprotein A-1, A-II and B, Complement C1q, C3 and C4, Plasminogen and Pre-albumin. The manufacturer's protocol was followed for the depletion procedure, briefly: 40 µl of plasma was diluted to 500 µl with Equilibration Buffer (EqB) and filtered at 0.22 µm with Corning Spin-X centrifuge tube filter. An immunodepletion spin column was equilibrated with EqB and 100 µl of filtered plasma was added and incubated for 20 min at room temperature. Unbound low abundant proteins were centrifuged at 4500 rpm for 1 min and the flow through was collected into a clean tube. Bound proteins were eluted off the column using elution buffer and the column re-equilibrated for the next depletion. Depleted plasma was concentrated using Vivaspin centrifugal concentrator and 100 µl of depleted-concentrated sample was added to the immunodepletion column for the second time to remove the high abundance proteins that did not bind during the first depletion. At the end of this stage, ~98% of high abundance proteins are depleted from the plasma.

Protein concentration was determined using Pierce Bicinchoninic Acid protein assay kit following the manufacturer's protocol (Thermo Scientific, Loughborough, UK).

Tryptic digestion

Low abundant proteins (typically 100µg) were chemically reduced by 100 mM dithiothreitol(final concentration 15mM) at 60°C for 15 minutes and alkylated in the dark with 200 mM iodoacetamide (final concentration 20mM) for 30 minutes at room temperature. Digestion of proteins into peptides was performed by adding trypsin at the concentration of 50:1 (protein:trypsin) for overnight incubation (approximately 16 hours) at 37°C. Digested proteins were concentrated for 90 minutes in SpeedVac followed by 3 hours freeze drying. Dried pellets were re-constituted in 10 µl of 0.1% formic acid + 10 µl of alcohol dehydrogenase (ADH) as an internal standard enabling absolute quantitation of the proteins post-analysis.

Fractionation

Tryptic peptides were separated on a RP column (XBridge BEH300 C18 1.0 x150mm) (Waters Ltd, Manchester, UK) in the first dimension by RPLC using a Varian ProStar 325 HPLC system (Agilent Technologies, Cheadle, UK). The HPLC system was comprised of an autosampler, a solvent delivery system and a UV detector (260nm). Solvent A: 200mM Ammonium Formate, pH 10 and solvent B: 10% 200mM Ammonium formate in Acetonitrile [14]. Flow rate was set at 0.1mL/min. At time zero A was 97% while B was 3%. B remained isocratic for 8min. B increased linearly to 40% at 63 min and to 88% B at 72 min and then returned to 3% at 80 min. The column was further equilibrated for 5 min at 3%. Total run time of 85 min. Fractions were collected every 6 min for 60 min. The final 25min was collected as a single fraction. Blanks were injected in between each analysis to eliminate carryover.

Targeted analysis

Individual samples from the 2 groups (Control and 30%) were immunodepleted , and subsequently underwent tryptic digestion, as described above. For targeted analysis, samples did not undergo fractionation prior to LC-IMS-MS/MS analysis.

2.3.3 Data Analysis

Raw data was then processed using Protein Lynx Global Server (PLGS) 3.0. PLGS 3.0 utilises the drift time of ion mobility-separated peptides to increase the specificity of alignment/association for precursor and product ions. Data were extracted, aligned and searched against the Uniprot human proteomic database, version 2013-11, appended with the alcohol dehydrogenase (*S. cerevisiae*) sequence. The ion accounting algorithm used has been described previously [15]. Expression^{E} was used to identify differences in protein expression between the samples using a Beyesian statistical model [13]. For targeted analysis, MassLynx 4.1 was used to extract and integrate the data. Integrated peaks were analysed using GraphPad Prism 6.02.

2.4 Metabolic Profiling

2.4.1 Instrumentation

The metabolomic analysis of mice plasma was performed on a Waters ACQUITY UPLC chromatograph with an in-built auto-sampler (Waters Corporation, Milford, MA) fitted with a Waters Acquity BEH (bridged ethylene hybrid) C_{18} 2.1 mm x 100 mm, 1.7 µm column (Waters Corporation, Milford, MA) maintained at 40°C. A VanGuard BEH C_{18} pre-column (Waters Corporation, Milford, MA) was fitted prior to the analytical column. Solvent A: HPLC grade water with 0.1% aqueous formic acid (v/v) and solvent B: acetonitrile with 0.1% formic acid (v/v). The flow rate was set to 0.5 mL/min. The mice plasma extract (10 µL injected) was analysed using the chromatographic gradient: 5% B (0-1 min), increased to 35% B (1-10 min). This was followed by a

column clean-up phase built into the method to reduce carry over and condition the column for analysis of subsequent samples, in which the mobile phase was increased to 95% B (10-11 min) and maintained at 95% B (11-12 min) before returning to initial conditions (13-15 min).

The LC system was coupled with a Waters Synapt HDMS ion-mobility mass spectrometer (Waters Corporation, Milford, MA) equipped with an electrospray interface. Electrospray ionisation conditions for the MS, with the ion source operated in positive ion mode were: capillary voltage 3.0 kV; cone voltage 30 V; source temperature 120°C; desolvation temperature 300°C; desolvation gas, N2 gas flow 600 L/hr; cone gas flow 30 L/hr. Data was acquired from 50 – 1200 Da with an inter-scan delay of 0.02 sec. The tri-wave (TWIMS) drift cell conditions were set at 30 mL/min drift gas (N2) with a variable travelling wave height ramp of 6.0 – 13.0 V and wave velocity of 300 m/s. The mass spectrum acquisition rate was 0.065 ms/scan. MassLynx v4.1 (Waters Corporation, Milford, MA) was used for controlling the setup and for data acquisition.

2.4.2 Sample Preparation

Plasma samples stored at -80°C were thawed at room temperature prior to sample clean-up by protein precipitation. Samples were analysed unpooled for metabolic profiling. Each thawed plasma sample was vortexed for 30 seconds followed by ultra-sonication for 1 minute to improve homogeneity. Protein precipitation was achieved by the addition of 200 µL acetonitrile to 100 µL of plasma [16]. The mixture was vortexed for 30 seconds followed by ultra-sonication for 1 minute and then subjected to centrifugation at 10,000 g for 10 minutes at ambient temperature. The supernatant was transferred to a fresh microcentrifuge tube and concentrated to near dryness by evaporation using a Turbovap LV concentration workstation (Caliper Life Sciences, Hopkinton, MA, USA), prior to being reconstituted in 50 µL water/acetonitrile (95:05) (v/v) with 0.1% formic acid. The

reconstituted plasma extract was placed in a 200 µL polypropylene micro-insert (Supelco, UK) in an autosampler vial (2 mL) for UPLC-MS analysis.

Quality Control

A quality control (QC) sample [17,18] was prepared by combining equal-volume sub-aliquots from all the samples in the study (Group A to D). The chromatographic column was conditioned at the start of the run by the injection of consecutive aliquots from five extracts derived from the QC samples. Randomisation of the sample list was carried out prior to analysis to remove bias. The QC sample was analysed after every 5 samples during the analytical runs. Blanks were analysed after every sample to ensure no-carryover between sample runs.

2.4.3 Data Analysis

The ion mobility dimension was collapsed to reduce the complexity of the data to retention time, *m/z* and intensity variables prior to data analysis. Data mining was initially carried out using MarkerLynx XS (Waters Corporation, Milford, MA). The parameters used for peak alignment were: retention time range 0-10 min, retention time tolerance was set to 0.1 min; mass range was limited to 50-500 Da and isotopic peaks were excluded from the analysis. The data matrix generated by the MarkerLynx software was normalised to the summed total ion intensity per chromatogram. The data were pareto-scaled prior to subjecting the data matrix to multivariate statistical analysis. Multivariate statistical analysis including PCA and PLS-DA was carried out using SIMCA-P + (Umetrics, Sweden). Ion mobility data was visualised and processed using Driftscope v2.1 (Waters Corporation, Milford, MA).

3.0 Results

3.1 Chemopreventive efficacy of dietary rice bran in *Apc*^{Min} mice

Mice received RiFiber with their diet at three concentrations (5, 15 and 30%) from week 4 of age and tumour number and burden were assessed in week 16, when most animals were moribund due to their intestinal adenoma load. Figure 1 shows that administration of RiFiber at the 30% dose decreased tumour burden and that tumour number was negatively correlated with rice bran dose. The highest dietary dose decreased the internal bleeding associated with adenoma load in the mice to a small but significant extent, as reflected by the haematocrit levels. These observations are fully consistent with a previous report on the effect of rice bran on adenoma development in this model [3]. There was no significant difference in body weight for any of the treatment groups at the time of culling (see Supplementary Figure 1).

3.2 Analysis of the plasma proteome in *Apc*^{Min} mice on rice bran

HDMS^{E} following immunodepletion of plasma has been shown to be reproducible in human samples [19] and provides accurate quantitation [20]. Therefore this approach was used to assess the effect of rice bran consumption on the mouse plasma proteome. Expression^{E} was used to establish significant differences ($p<0.05$) in protein expression which were observed between control mice and mice on 30% rice bran (Table 3). Thirty eight proteins were differentially expressed (36 up-regulated, 2 down-regulated in the intervention group). A literature search (Pubmed, Web of Science) suggested that some of these proteins are associated with carcinogenesis. Three molecules;

gelsolin, adiponectin and α-2HS glycoprotein were selected for further interrogation. Five mass spectrometric transitions for each of these three peptides were derived from the discovery data to help verify their presence in individual plasma samples using single LC-MS-SRM analysis. Figure 2 demonstrates agreement between discovery and individual sample analyses. Expression of the three proteins namely Gelsolin (2.5 fold; p = 0.018), α-2HS (2.13 fold; p = 0.013) and adiponectin (4.5 fold; p = 0.0003) in the individual samples was significantly higher in the plasma of mice on rice bran than in control mice. These values compare with fold differences in the discovery analyses of 1.77, 1.14 and 1.58, respectively. Correlation analysis between the number of adenomas and peak area did not correlate for Gelsolin (r^2=0.06 p=0.37) and α-2HS glycoprotein (r^2=0.01 p=0.98) but a weak correlation (r^2=0.22 p=0.07) was observed for adiponectin (Supplementary Figure 2).

3.4 Plasma Metabolic Profiling in *Apc^{Min}* mice on rice bran.

The LC-MS data derived from the metabolic profiling was mined using MarkerLynx XS (Waters Corporation) and processed using multivariate statistical analysis including principal component analysis (PCA) and orthogonal partial least squares discriminant analysis (OPLS-DA). Supplementary Figure 2 shows the PCA plot of all the samples from the study (Group A – control mice to Group D – mice on 30% RiFiber) and the quality control (QC) samples. The tight clustering of the QC samples towards the middle of the PCA plot (circled in supplementary Figure 3) indicates that data quality is fit for purpose for metabolite profiling studies. The numbered QC samples did not show any particular trend, their distribution was governed by random order within the cluster of samples (data not shown) indicative of absence of drift in system performance, which would bias the analysis. The PCA plot (Figure 3A) shows a partial separation of metabolic profiles in the plasma between control

mice (Group A) and mice on 30% rice bran (Group D). An OPLS-DA model based on K=315 variables produced a clear separation between these two groups. Plasma samples from control mice showed a wider within-sample-group variation than those from mice on 30% rice bran, as reflected by the spread of samples from control mice samples across principal component 2 in comparison to those from mice on rice bran which were clustered more tightly (Figure 3B). The control group consisted of 4 male and 3 female mice whereas the 30% rice bran group contained 4 male and 4 female mice. The OPLS-DA plot does not suggest sub-clustering of samples indicating that the effect due to rice bran is independent of gender.

A loadings S-plot (Supplementary Figure 4) was constructed using the OPLS-DA model shown (Figure 3b) to segregate potential metabolite ions and allow visualisation of the distribution of the data from the OPLS-DA model for the two classes. The S-plot indicates the presence of seven metabolite ions, five up-regulated and two down-regulated expressed at different levels between control mice and mice on 30% rice bran (Table 2). These metabolites were interrogated *via* their extracted ion chromatograms (EIC) as shown in Figure 4 for the *m/z* 212 ion. Trends across groups for individual metabolites reflect differences in abundance between groups illustrated in Figure 4a and 4b for this up-regulated ion ($p = 0.0003$) . This data suggests that consumption of RiFiber at 30% in the diet affects circulating metabolites, although conclusive identification of up- and down- regulated metabolic species requires further work.

4.0 Discussion

The chemoprevention of CRC by fibre is an attractive dietary strategy for the reduction of incidence of CRC in individuals susceptible to this disease. A recent paper suggests that once weekly consumption of brown rice can reduce the risk of colonic polyp formation by 40% [21]. This and other results obtained with high fibre-containing dietary sources render fibre consumption

potentially worthy of inclusion into public disease prevention strategies. However, any extensive propagation of fibre consumption as a CRC prevention strategy necessitates a thorough understanding of the biochemical mechanisms involved. Animal models allow the study of drivers of carcinogenesis which can be confounded by chemopreventive diets, and they can help formulate hypotheses pertinent to chemopreventive efficacy under strictly defined experimental conditions. The *Apc*^{Min} mouse is such a model permitting exploration of aspects of the molecular development of colorectal adenomas which often precede full-blown colorectal carcinomas [4]. The model mirrors the familial adenomatous polyposis (FAP) syndrome in humans in whom deletions in the *APC* gene result in an almost 100% lifetime risk of CRC development. Whilst *Apc*^{Min} mice develop adenomas in their small intestine, neoplasms in FAP sufferers tend to be localised in the colon [22]. In the work presented above an attempt was made to detect plasma molecules indicative of anti-adenomatous events in mice in which adenoma development was significantly confounded by consumption of rice bran. The results clearly indicate that the dietary intervention affects circulating protein and metabolite patterns.

PCA of the metabolic profiling data showed partial resolution of plasma samples between control mice and mice on 30% RiFiber, although at lower doses such a separation was not observed. OPLS-DA was used to identify molecules potentially responsible for the separation. It is conceivable that some of these identified metabolites are pure markers of RiFiber consumption; however, further work is required for definitive structural allocation of these metabolites. These metabolic changes conceivably reflect the influence of RiFiber consumption on microbial events linked to adenomagenesis [23]. Modulation of gut microflora is increasingly seen as a potential mediator of colorectal carcinogenesis and may act as a molecular signature [24]. Such signatures have been observed before in human blood [25].

Three of the molecular changes detected here by the proteomic analysis were further tested in targeted analyses. Adiponectin, gelsolin and α-2HSGP were found to be increased in the plasma of

mice which consumed RiFiber as compared to controls. Elevated levels of adiponectin have previously been shown to accompany suppression of carcinogenesis in *Apc*^{Min} mice [26]. Adiponectin has both anti-inflammatory and insulin-sensitising effects [27]. These effects have been hypothesised as key mechanisms that associate obesity with colorectal carcinogenesis [28]. In this study a change in body weight was not observed for any of the treatment groups indicating independence from potential obesity confounders. A recent EPIC study found that total circulating and non-high molecular weight adiponectin was inversely associated with CRC [29]. The mechanism underlying this relationship is unclear, but it has been suggested that the AMPK/mTOR pathway may be implicated [30].

Gelsolin has been detected as a biomarker previously for cancers other than CRC, and in one study plasma gelsolin was a biomarker of distant organ metastasis [31]. α-2HSGP has been suggested to be a potential marker of CRC [32]. Paradoxically in the study described here, circulating levels of both proteins were up-, not down-regulated, inconsistent with the cancer chemopreventive activity of RiFiber. Nevertheless the data demonstrates that changes occur in the plasma at the molecular level which probably emanate from intestinal adenomas. These changes might provide potential biomarkers of efficacy in future chemopreventation trials of rice bran in individuals at risk of colorectal carcinogenesis.

The dose of rice bran ingested by the mice in the 30% bran group, in which the described molecular changes were observed, was about 15 g per kg mouse, which equates to approximately 1.2g/day per kg in humans, i.e. ~85 g in a 70 kg human, when calculated using Freireich rodent–to human extrapolation [33]. This is a rather hefty - but still just feasible - amount in the light of the recommended daily fibre intake in a human which is about 18g per day [34]. More preclinical work and large sample numbers are required to confirm the relevance of the observations proffered here [35].

Acknowledgments: The authors would like to thanks the HOPE against Cancer Foundation for funding this study. We also thank RiceBran Technologies (formerly Nutracea, Scottsdale, AR, USA) for generous support of the study. Work was also supported by a Cancer Research UK Programme grant (C325/A13101).

Author contributions: Study design and experimental plan (CC, DJ, AG and SS). Sample preparation (AM, EHG, LN). Sample and data analysis (AM, LN, AH). Manuscript Preparation (DJ, CC, AG, AM and SS).

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

References

- [1] Norat T, Aune D, Chan D, Romaguera D, Fruits and vegetables: updating the epidemiologic evidence for the WCRF/AICR lifestyle recommendations for cancer prevention. *Cancer Treat Res.* 2014, *159*, 35-50.
- [2] Cai H, Al-Fayez M, Tunstall RG, Platton S, Greaves P et al., The rice bran constituent tricin potently inhibits cyclooxygenase enzymes and interferes with intestinal carcinogenesis in ApcMin mice. *Mol Cancer Ther.* 2005, *4*, 1287-92.
- [3] Verschoyle RD, Greaves P, Cai H, Edwards RE, Steward WP et al., Evaluation of the cancer chemopreventive efficacy of rice bran in genetic mouse models of breast, prostate and intestinal carcinogenesis. *Br J Cancer.* 2007, *96*, 248-54.
- [4] Moser AR, Luongo C, Gould KA, McNeley MK, Shoemaker AR et al., ApcMin: a mouse model for intestinal and mammary tumorigenesis. *Eur J Cancer.* 1995, *31*, 1061-4.
- [5] Steward WP, Brown K. Cancer chemoprevention: a rapidly evolving field. *Br J Cancer.* 2013, *109* ,1-7.
- [6] Rahman MM, Seo YR. Discovery of potential targets of selenomethionine-mediated chemoprevention in colorectal carcinoma mouse model using proteomics analysis. *Carcinogenesis*. 2013, *34*, 1575-84.
- [7] Vyas AR, Hahm ER, Arlotti JA, Watkins S, Stolz DB et al., Chemoprevention of prostate cancer by d,l-sulforaphane is augmented by pharmacological inhibition of autophagy. *Cancer Res.* 2013, *73*, 5985-95.
- [8] Ma YC, Li C, Gao F, Xu Y, Jiang ZB et al., Epigallocatechin gallate inhibits the growth of human lung cancer by directly targeting the EGFR signaling pathway. *Oncol Rep.* 2014, 31, 1343-9.
- [9] George J, Singh M, Srivastava AK, Bhui K, Roy P et al., Resveratrol and black tea polyphenol combination synergistically suppress mouse skin tumors growth by inhibition of activated MAPKs and p53. *PLoS One.* 2011, *6*, e23395.
- [10] Backshall A, Alferez D, Teichert F, Wilson ID, Wilkinson RW et al., Detection of metabolic alterations in non-tumor gastrointestinal tissue of the Apc(Min/+) mouse by (1)H MAS NMR spectroscopy. *J Proteome Res.* 2009, *8*, 1423-30.
- [11] http://www.testdiet.com/cs/groups/lolweb/@testdiet/documents/web_content/mdrf/mdi2/~edisp/ducm04_026410.pdf
- [12] Beer LA, Wang H, Tang HY, Cao Z, Chang-Wong T, Tanyi JL, Zhang R, Liu Q, Speicher DW. Identification of multiple novel protein biomarkers shed by human serous ovarian tumors into the blood of immunocompromised mice and verified in patient sera. *PLoS One*. 2013;8(3):e60129.
- [13] Silva JC, Denny R, Dorschel CA, Gorenstein M, Kass IJ, et al., Quantitative proteomic analysis by accurate mass retention time pairs. *Anal Chem.* 2005, *77*, 2187-200.
- [14] Gilar M, Olivova P, Daly AE, Gebler JC. Two-dimensional separation of peptides using RP-RP-HPLC system with different pH in first and second separation dimensions. *J Sep Sci.* 2005, 28, 1694-703.

- [15] Li, G, Vissers, JPC, Silva, JC, Golick, D, Gorenstein, MV et al., Database searching and accounting of multiplexed precursor and product ion spectra from the data independent analysis of simple and complex peptide mixtures. *Proteomics*. 2009, *9*, 1696-1719.

- [16] Polson C, Sarkar P, Incledon B, Raguvaran V, Grant R. Optimization of protein precipitation based upon effectiveness of protein removal and ionization effect in liquid chromatography-tandem mass spectrometry. *Journal of chromatography. B, Analytical technologies in the biomedical and life sciences*, 2003, *785*, 263–75.

- [17] Want EJ, Wilson ID, Gika H, Theodoridis G, Plumb RS et al., Global metabolic profiling procedures for urine using UPLC-MS. *Nat Protoc.,* 2010, *5*, 1005-18.

- [18] Want EJ, Masson P, Michopoulos F, Wilson ID, Theodoridis G et al., Global metabolic profiling of animal and human tissues via UPLC-MS. *Nat Protoc.* 2013, 8, 17-32

- [19] Hakimi A, Auluck J, Jones GD, Ng LL, Jones DJ. Assessment of reproducibility in depletion and enrichment workflows for plasma proteomics using label-free quantitative data-independent LC-MS. *Proteomics*. 2014, *14*, 4-13.

- [20] Daly CE, Ng LL, Hakimi A, Willingale R, Jones DJ. Qualitative and quantitative characterization of plasma proteins when incorporating traveling wave ion mobility into a liquid chromatography-mass spectrometry workflow for biomarker discovery: use of product ion quantitation as an alternative data analysis tool for label free quantitation. *Anal Chem.* 2014, 86, 1972-9.

- [21] Tantamango YM, Knutsen SF, Beeson L, Fraser G, Sabate J. Association between dietary fiber and incident cases of colon polyps: the adventist health study. *Gastrointest Cancer Res.* 2011, *4*, 161-7.

- [22] Tammariello AE, and Milner JA. Mouse Models for Unravelling the Importance of Diet in Colon Cancer Prevention. *J Nutr Biochem.2010*, *21*, 77–88.

- [23]Henderson AJ, Ollila CA, Kumar A, Borresen EC, Raina K, Chemopreventive Properties of Dietary Rice Bran: Current Status and Future Prospects. *Adv Nutr.,* 2012, 3, 643-653.

- [24] Holmes E, Li JV, Marchesi JR, Nicholson JK. Gut microbiota composition and activity in relation to host metabolic phenotype and disease risk. *Cell Metab,.* 2012, 16, 559-64.

- [25] Tan B, Qiu Y, Zou X, Chen T, Xie G et al., Metabonomics identifies serum metabolite markers of colorectal cancer. *J Proteome Res.* 2013, 12, 3000-9.

- [26] Otani K, Kitayama J, Yasuda K, Nio Y, Iwabu M et al. Adiponectin suppresses tumorigenesis in Apc(Min)(/+) mice. *Cancer Lett.* 2010 ,*288*, 177-82.

- [27]Tilg H, Moschen AR. Adipocytokines: mediators linking adipose tissue, inflammation and immunity. *Nat Rev Immunol.* 2006, *6*, 772-83.

- [28] Song M, Zhang X, Wu K, Ogino S, Fuchs CS et al., Plasma adiponectin and soluble leptin receptor and risk of colorectal cancer: a prospective study. *Cancer Prev Res (Phila)*. 2013, *6*, 875-85.

- [29] Aleksandrova K, Boeing H, Jenab M, Bueno-de-Mesquita HB, Jansen E et al.,, Total and high-molecular weight adiponectin and risk of colorectal cancer: the European Prospective Investigation into Cancer and Nutrition Study. *Carcinogenesis.* 2012, 33, 1211-8.

- [30] Sugiyama M, Takahashi H, Hosono K, Endo H, Kato S et al.,, Adiponectin inhibits colorectal cancer cell growth through the AMPK/mTOR pathway. *Int J Oncol*. 2009, 34, 339-44.

- [31]Tsai MH, Wu CC, Peng PH, Liang Y, Hsiao YC et al., Identification of secretory gelsolin as a plasma biomarker associated with distant organ metastasis of colorectal cancer. *J Mol Med (Berl).* 2012 , *90*, 187-200.

- [32] Fan NJ, Kang R, Ge XY, Li M, Liu Y et al., Identification alpha-2-HS-glycoprotein precursor and tubulin beta chain as serology diagnosis biomarker of colorectal cancer. *Diag Pathol.* 2014 *12*, 9-53.

- [33] Freireich EJ, Gehan EA , Rall DP, Schmidt LH, Skipper HE. Quantitative comparison of toxicity of anticancer agents in mouse, rat, dog, monkey and man. Cancer Chemother Rep.1966;50(4):219-244.

- [34] http://www.nutrition.org.uk/nutritionscience/nutrients/dietary-fibre.html

- [35] Forster GM, Raina K, Kumar A, Kumar S, Agarwal R et al., Rice varietal differences in bioactive bran components for inhibition of colorectal cancer cell growth. Food Chem. 2013, 141,,1545-52.

Figure Legends

Figure 1. Effect of lifetime dietary fibre administration on (i) adenoma burden, (ii) adenoma number and (iii) haematocrit value of ApcMin mice. Values are the mean ± SD of between 19-22 animals per group. Where data point is significantly different to the control p values are shown.

Figure 2. Verification of discovery phase with label free LC-MS-SRM analysis for Gelsolin, Alpha 2 HS Glycoprotein and Adiponectin in individual samples. All three proteins demonstrate statistically significant differences (p = 0.018, 0.013 and 0.0003 respectively) between 0 and 30% RiFiber.

Figure 3. PCA and OPLS-DA analysis of 0 (Group A) and 30% (Group D) RiFiber. The multivariate analysis allows observable changes in the samples to be reflected in differentiation between clusters. A clear separation can be seen in the OPLS-DA analysis (Figure 3B) for the 2 groups.

Figure 4. (A) Representive overlaid UHPLC-MS extracted ion chromatogram (EIC) of m/z 212.95 for Zero (A) and 30% (D) RiFiber, one of the ions identified by the O-PLSDA model's loading s-plot for being responsible for separation between 0 and 30% RiFiber (p=0.0003). (B) The EIC intensities for *m/z* 212.95 showing differences between 0 and 30% RiFiber plasma samples.

Supplementary Figure 1. Effect of rice bran on whole body weight of male and female Apc^{Min} mice. Mice received control diet or diet fortified with RiFiber at 5, 15 or 30%.

Supplementary Figure 2. Correlation plots of adenoma number vs peak area is plotted for Gelsolin, α-2HS glycoprotein and adiponectin. There is no correlation observed for Gelsolin (r^2 = 0.06, p = 0.37) and α-2HS glycoprotein (r^2=0.01, p = 0.98) for adenoma number and peak area, however there is a weak correlation (r^2=0.22, p=0.07) for adiponectin.

Supplementary Figure 3. PCA plot showing the separation between all four groups. A = control (0), B = 5%, C=10% and D= 30% RiFiber administration. QC are the QC samples which are pooled form all samples. QCs clustering demonstrates that the platform performance is consistent throughout the study analysis. However, there is little clustering between the four studied groups.

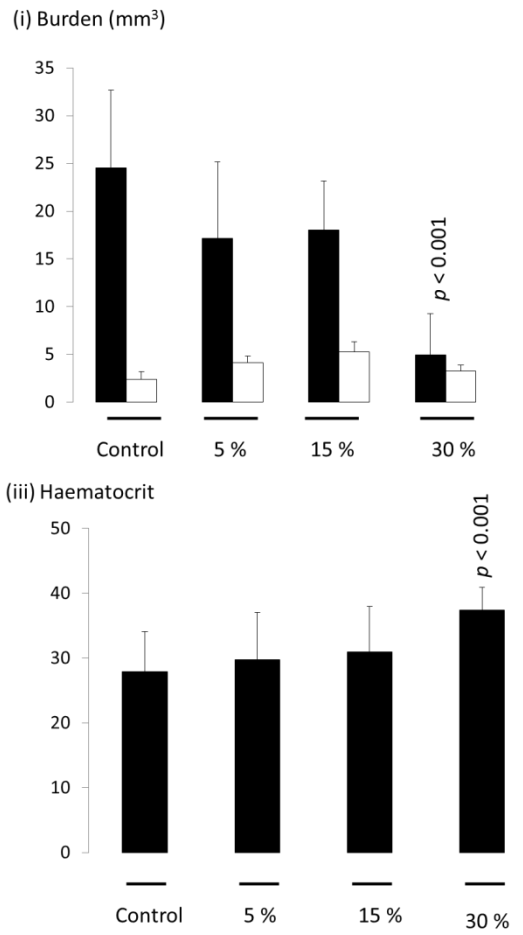
Supplementary Figure 4. S-Plot analysis of all features observed between 0 and 30% RiFiber. The features showing the largest variance are potential biomarkers and are circled in red.

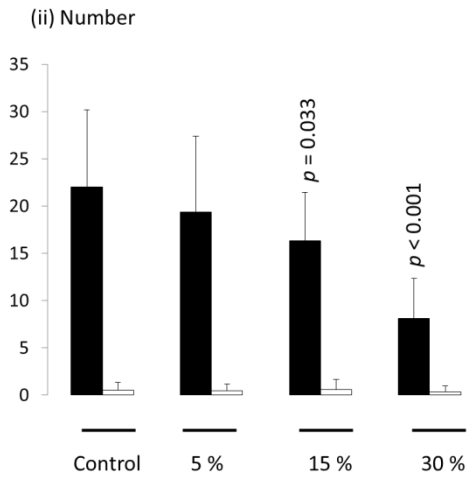
Table 1. Numbers of APC^{Min} mice used for the different groupings with gender information.

Table 2. Identified ions from metabonomic analysis that provide most separation in PLSDA analysis. Retention time and mass (Da) information is supplied.

Table 3. Selection of proteins with observable expression changes following proteomic analysis.

Figure 1





Black bars = Small Intestine
White Bars = Colon

Figure 2

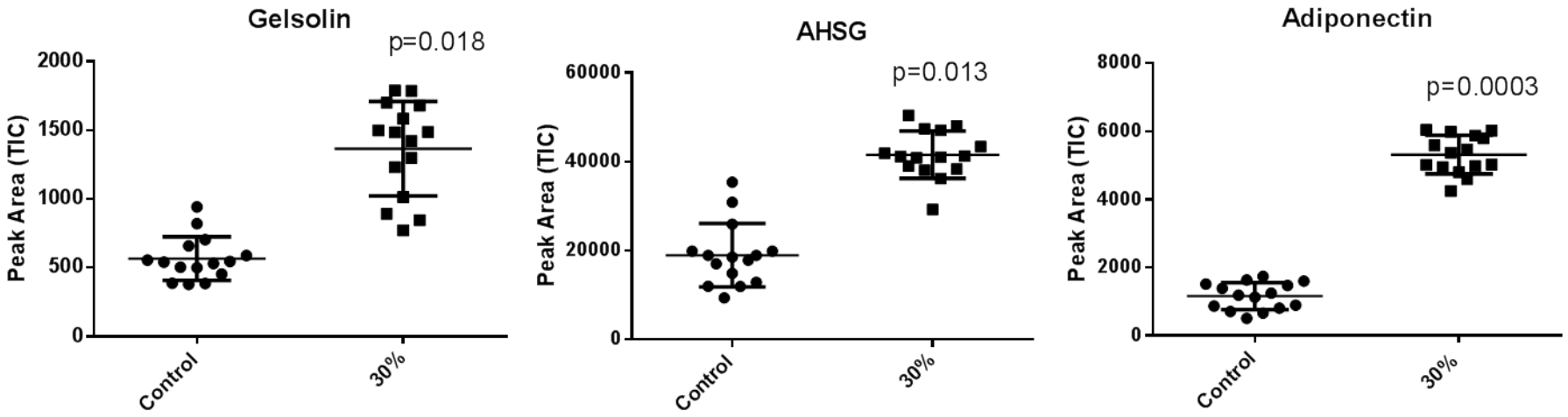
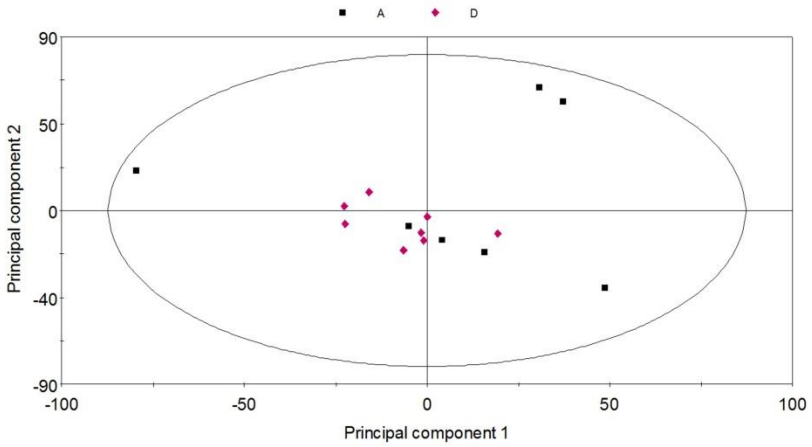

Gelsolin
p=0.018
Peak Area (TIC)
2000
1500
1000
500
0
Control
30%
AHSG
p=0.013
Peak Area (TIC)
60000
40000
20000
0
Control
30%
Adiponectin
p=0.0003
Peak Area (TIC)
8000
6000
4000
2000
0
Control
30%

Table 1

Group	Supplemented Bran (%)	Total (n)	Male	Female
Group A	0	7	4	3
Group B	5	4	4	0
Group C	15	16	9	7
Group D	30	8	4	4
Total		***35***	*21*	*14*

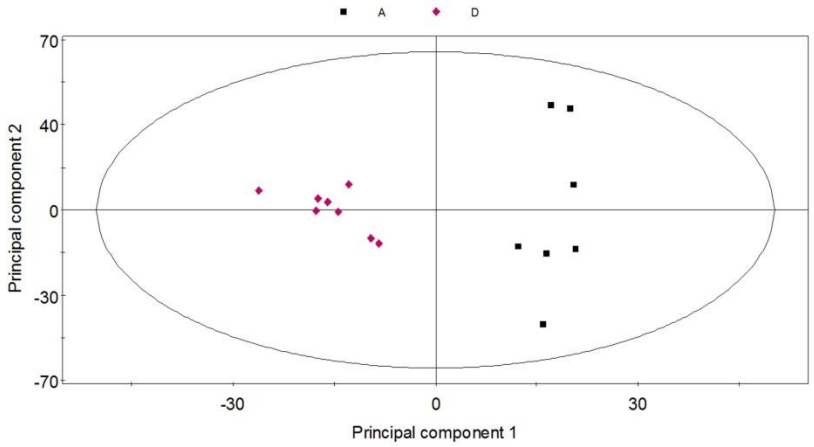
Figure 3

A



PCA

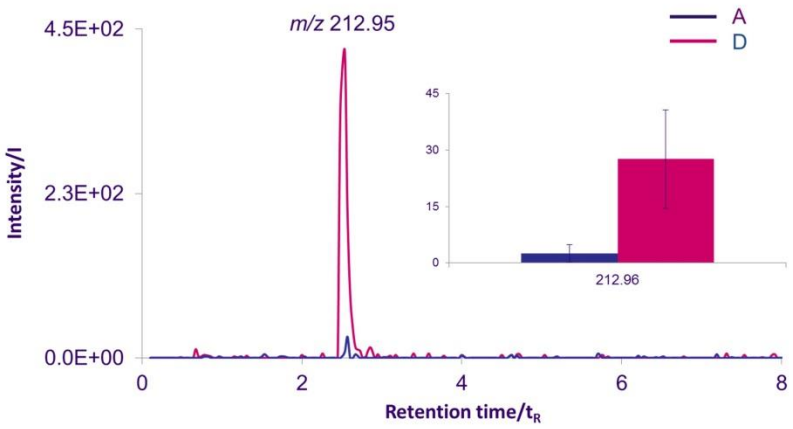
B



OPLS-DA

Figure 4

A


m/z 212.95
A
D
4.5E+02
2.3E+02
0.0E+00
Intensity/I
0
2
4
6
8
Retention time/t_R
45
30
15
0
212.96

B

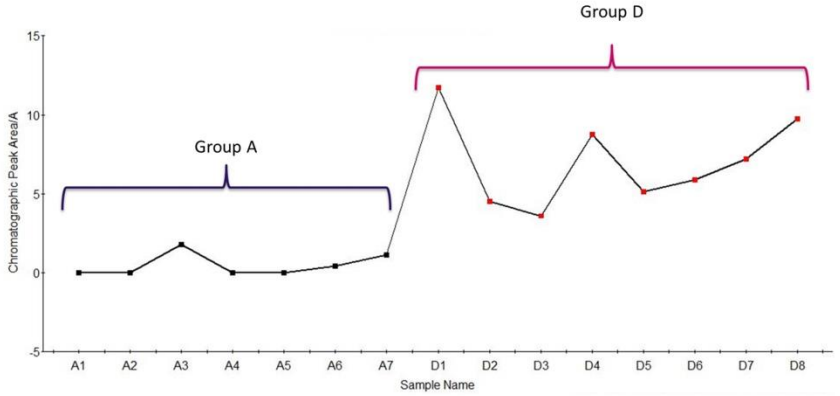

Group D
Group A
15
10
5
0
-5
Chromatographic Peak Area/A
A1
A2
A3
A4
A5
A6
A7
D1
D2
D3
D4
D5
D6
D7
D8
Sample Name

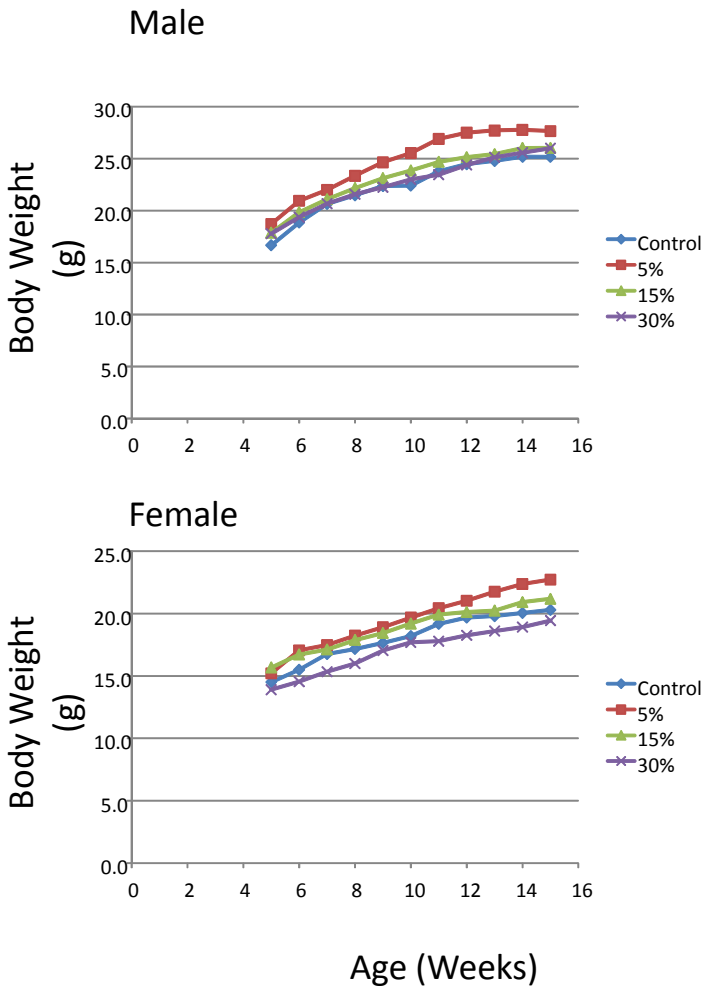
Table 2

No.	Retention Time (min)	*m/z* ±0.02	Regulation
1	2.52	148.99	Up-regulated
2	7.71	189.02	Up-regulated
3	2.54	212.95	Up-regulated
4	3.19	218.01	Up-regulated
5	2.54	218.97	Up-regulated
6	0.92	160.07	Down-regulated
7	4.48	226.97	Down-regulated

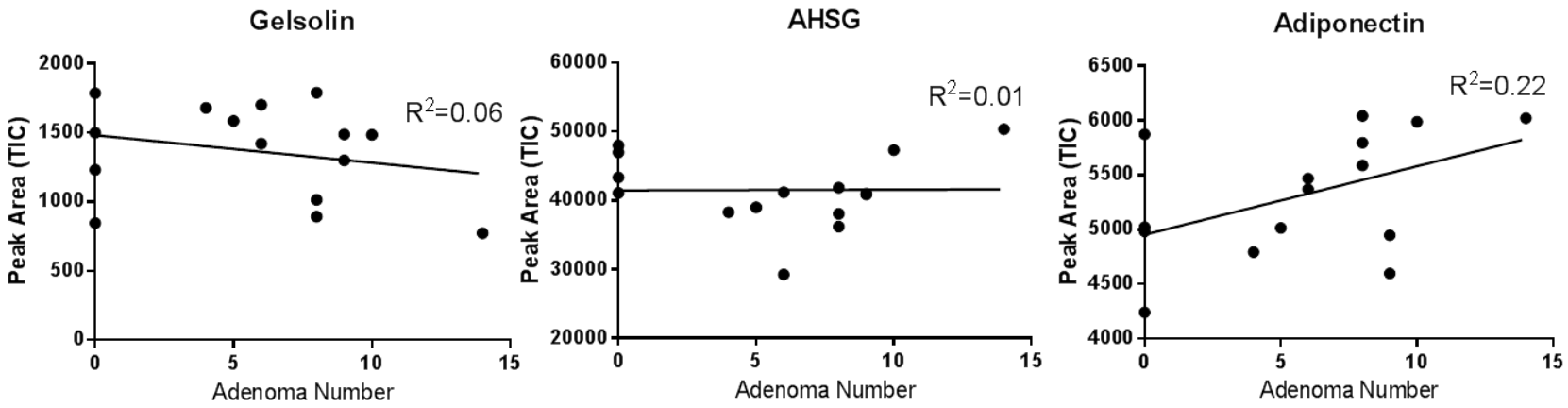
Table 3

Protein	Fold Change	Direction
Serine protease inhibitor A3G	4.35	↑
Haptoglobin	3.35	↑
Major urinary protein 2	3	↑
Fibrinogen gamma chain	2.83	↑
Fibrinogen beta chain	2.77	↑
Alpha 1 acid glycoprotein 1	2.44	↑
Ig mu chain C region secreted form	2.39	↑
Ig kappa chain C region	2.27	↑
Alpha 2 macroglobulin	2.05	↑
Apolipoprotein A 1	1.95	↑
Serine protease inhibitor A3 K	1.92	↑
Serum albumin	1.82	↑
Immunoglobulin J chain	1.77	↑
Gelsolin	1.77	↑
Ig gamma 1 chain C region secreted form	1.7	↑
Serum amyloid P component	1.65	↑
Alpha 1 antitrypsin 1 4	1.6	↑
Adiponectin	1.58	↑
Interleukin 1 receptor accessory protein	1.57	↑
Complement factor H	1.52	↑
Alpha 1 antitrypsin 1 1	1.52	↑
Kininogen 1	1.49	↑
Complement factor 1	1.49	↑
Fetuin B	1.4	↑
Angiotensinogen	1.38	↑
Hemopexin	1.34	↑
Alpha 1 antitrypsin 1 2	1.3	↑
Beta 2 glycoprotein 1	1.28	↑
Apolipoprotein 1 IV	1.28	↑
Liver carboxylesterase N	1.27	↑
Murinoglobulin 1	1.27	↑
Afamin	1.25	↑
Alpha 1 antitrypsin 1 5	1.2	↑
Alpha 2 HS glycoprotein	1.14	↑
Complement C3	1.08	↑
Serotransferrin	1.06	↑
Hemoglobin subunit epsilon Y2	0.25	↓
Carboxylesterase 3	0.22	↓

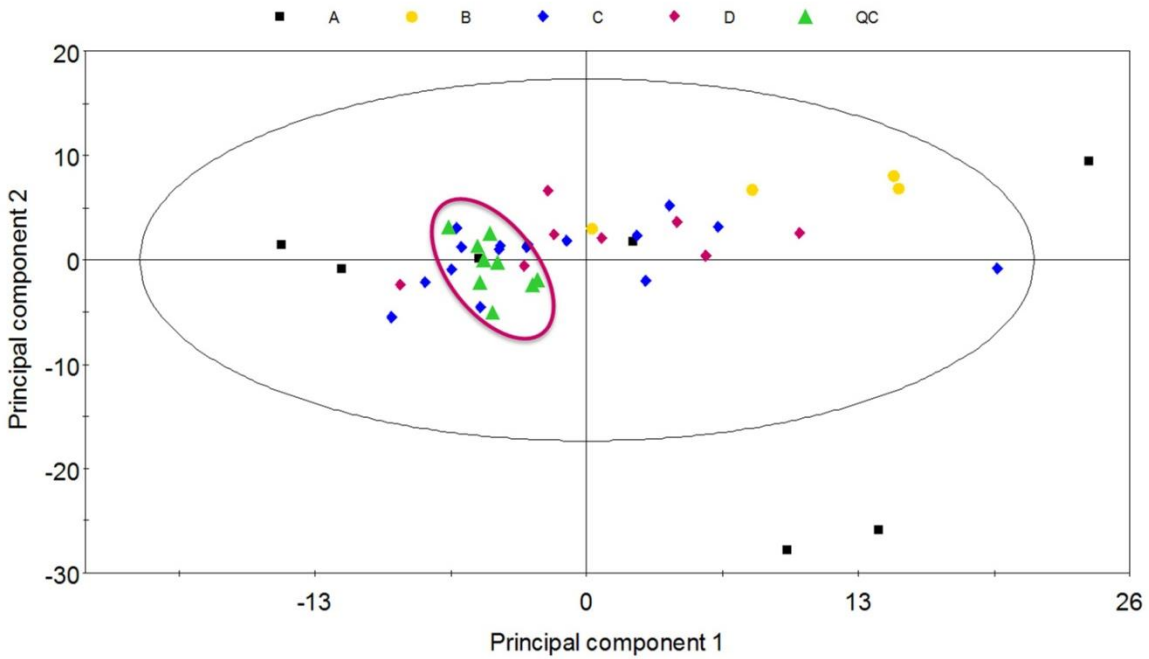
Supplementary Figure 1. Effect of rice bran on whole body weight of male and female ApcMin mice. Mice received control diet or diet fortified with RiFiber at 5, 15 or 30%.



Supplementary Figure 2. Correlation plots of adenoma number vs peak area is plotted for Gelsolin, α-2HS glycoprotein and adiponectin. There is no correlation observed for Gelsolin (r^2 = 0.06, p = 0.37) and α-2HS glycoprotein (r^2=0.01, p = 0.98) for adenoma number and peak area, however there is a weak correlation (r^2=0.22, p=0.07) for adiponectin.



Supplementary Figure 3. PCA plot showing the separation between all four groups. A = control (0), B = 5%, C=10% and D= 30% RiFiber administration. QC are the QC samples which are pooled form all samples. QCs clustering demonstrates that the platform performance is consistent throughout the study analysis. However, there is little clustering between the four studied groups.



Supplementary Figure 4. S-Plot analysis of all features observed between 0 and 30% RiFiber. The features showing the largest variance are potential biomarkers and are circled in red.