Genome Medicine

CrossMark

The influence of a short-term gluten-free diet on the human gut microbiome

Marc Jan Bonder[1†], Ettje F. Tigchelaar[1,2†], Xianghang Cai[3†], Gosia Trynka[4], Maria C. Cenit[1], Barbara Hrdlickova[1], Huanzi Zhong[3], Tommi Vatanen[5,6], Dirk Gevers[5], Cisca Wijmenga[1,2], Yang Wang[3†] and Alexandra Zhernakova[1,2*†]

Abstract

Background: A gluten-free diet (GFD) is the most commonly adopted special diet worldwide. It is an effective treatment for coeliac disease and is also often followed by individuals to alleviate gastrointestinal complaints. It is known there is an important link between diet and the gut microbiome, but it is largely unknown how a switch to a GFD affects the human gut microbiome.

Methods: We studied changes in the gut microbiomes of 21 healthy volunteers who followed a GFD for four weeks. We collected nine stool samples from each participant: one at baseline, four during the GFD period, and four when they returned to their habitual diet (HD), making a total of 189 samples. We determined microbiome profiles using 16S rRNA sequencing and then processed the samples for taxonomic and imputed functional composition. Additionally, in all 189 samples, six gut health-related biomarkers were measured.

Results: Inter-individual variation in the gut microbiota remained stable during this short-term GFD intervention. A number of taxon-specific differences were seen during the GFD: the most striking shift was seen for the family *Veillonellaceae* (class *Clostridia*), which was significantly reduced during the intervention ($p = 2.81 \times 10^{-05}$). Seven other taxa also showed significant changes; the majority of them are known to play a role in starch metabolism. We saw stronger differences in pathway activities: 21 predicted pathway activity scores showed significant association to the change in diet. We observed strong relations between the predicted activity of pathways and biomarker measurements.

Conclusions: A GFD changes the gut microbiome composition and alters the activity of microbial pathways.

Keywords: Microbiome, Gluten-free diet, Biomarker, Observation study

Background

Gluten is a major dietary component of wheat, barley, and rye. In genetically susceptible individuals, the consumption of gluten triggers the development of coeliac disease – an autoimmune disorder commonly seen in populations of European ancestry (with a frequency of approximately 1 %) [1]. In the absence of any medication, the only treatment is a life-long gluten-free diet (GFD), which is effective and well tolerated by the majority of patients. Non-coeliac gluten sensitivity, another common disorder linked to the consumption of gluten-containing food and resulting in a range of symptoms of intestinal discomfort (such as diarrhea and abdominal pain), has also been shown to improve on a GFD [2, 3]. More recently, a GFD is being considered as a way to ameliorate symptoms in patients with irritable bowel syndrome (IBS) [4].

However, beyond these medical indications, more and more individuals are starting on a GFD to improve their health and/or to control weight. The diet's popularity has risen rapidly in the last few years, making it one of the most popular diets worldwide, along with a low-carbohydrate diet and a fat-free diet. The numbers of those adopting the diet for non-medical reasons now surpass the numbers of those who are addressing a permanent gluten-related disorder [3].

Several studies have reported the effect of a GFD on the composition of the gut microbiome in coeliac disease

* Correspondence: sashazhernakova@gmail.com
†Equal contributors
[1]Department of Genetics, University of Groningen, University Medical Centre Groningen, Groningen, The Netherlands
[2]Top Institute Food and Nutrition, Wageningen, The Netherlands
Full list of author information is available at the end of the article

Bonder *et al. Genome Medicine* (2016) 8:45

Page 2 of 11

patients [5–7]. In these studies, the microbiome composition in coeliac patients on a GFD was compared with untreated patients and healthy individuals. The most consistent observation across these studies is the difference in the abundance and diversity of *Lactobacillus* and *Bifidobacterium* in the treated and untreated coeliac disease patients. It should be noted that these studies were relatively small (seven to 30 participants in each group). Specifically, De Palma et al. [8] assessed the effect of a one-month GFD on ten healthy individuals, but the study was limited to the use of non-sequence based methods, including FISH and qPCR. Their study described how *Bifidobacterium*, *Clostridium lituseburense*, *Faecalibacterium prausnitzii*, *Lactobacillus*, and *Bifidobacterium longum* were decreased during GFD, whereas *Escherichia coli*, *Enterobacteriaceae*, and *Bifidobacterium angulatum* were increased. To the best of our knowledge, there has been no comprehensive analysis of the effect of a GFD on the entire gut microbiome composition using a next-generation sequencing approach.

The effect of other diet interventions on the microbiome composition was recently studied using the 16S rRNA sequencing method [9]. In particular, it was shown that a short-term animal-based diet led to an increased abundance of bile-tolerant microorganisms (*Alistipes*, *Bilophila*, and *Bacteroides*) and a decreased abundance of Firmicutes, which metabolize dietary plant polysaccharides (*Roseburia*, *Eubacterium rectale*, and *Ruminococcus bromii*) [9].

In this work we assessed the effect of a GFD on gut microbiota using the next-generation 16S rRNA sequencing method. The analysis was performed in 189 samples, representing up to nine time points for 21 individuals. We investigated the diet-related changes both on the level of taxonomic units as well as on the predicted bacterial pathways. Next to this, we assessed a set of selected biomarkers to assess the gut health in relation to changes in bacterial composition and their association to a GFD. Our study offers insights into the interaction between the gut microbiota and a GFD.

Methods

Study design

We enrolled 21 participants (nine men and twelve women), without any known food intolerance and without known gastrointestinal disorders, in our GFD study for 13 weeks (Fig. 1). After baseline measurements (T = 0), all the participants started a GFD for four weeks (T = 1–4), followed by a "wash-out" period of five weeks. Subsequently, data were collected when they returned to their habitual diets (HD, gluten-containing) for a period of four weeks (T = 5–8) (Fig. 1). Fecal samples were collected at all time points. Blood was collected at baseline, at T = 2 and T = 4 on GFD, and at T = 6 and T = 8 on HD.
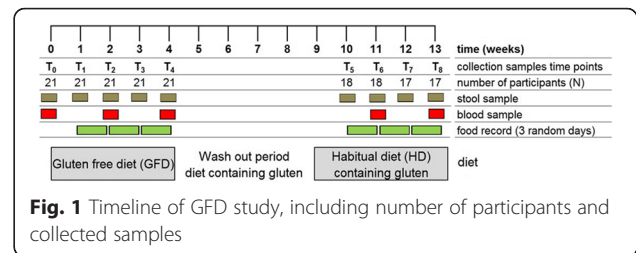

Fig. 1 Timeline of GFD study, including number of participants and collected samples

The participants were aged between 16 and 61 years (mean age, 36.3 years). Mean BMI was 24.0 and 28.6 % (n = 6) of participants were smokers. The majority of participants were European (n = 19), two participants were South American, and one was Asian. Except for one, none of the participants had taken an antibiotic treatment for the year prior to the study start. In both diet periods (GFD, HD), participants kept a detailed three-day food record. All 21 participants completed the GFD period; for 17 participants all data points were available. An overview of the participants' characteristics can be found in Additional file 1: Figure S1.

Written consent was obtained from all participants and the study followed the sampling protocol of the LifeLines-DEEP study [10], which was approved by the ethics committee of the University Medical Centre Groningen, document no. METC UMCG LLDEEP: M12.113965.

Gluten-free diet and dietary intake assessment

Methods to assess GFD adherence and dietary intake have been described previously by Baranska et al. [11] In short, before the start of the study, the participants were given information on gluten-containing food products by a dietician and they were instructed how to keep a three-day food record. The food records were checked for completeness and the macronutrient intake was calculated. Days on which a participant had a daily energy intake below 500 kcal or above 5000 kcal were excluded from our analysis (n = 2). Of 21 participants, 15 (71 %) completed the dietary assessments; three were excluded from food intake analysis because of incomplete food records. We used the paired t-test to compare group means between GFD and HD.

Blood sample collection

Participants' blood samples were collected after an overnight fast by a trained physician assistant. We collected two EDTA tubes of whole blood at baseline (T0) and during the GFD period at time points T2 and T4; during the HD period one EDTA tube was collected at time points T6 and T8. Plasma was extracted from the whole blood within 8 h of collection and stored at –80 °C for later analysis.

Bonder *et al. Genome Medicine* (2016) 8:45

Page 3 of 11

Microbiome analysis

Fecal sample collection

Fecal samples were collected at home and immediately stored at –20 °C. At the end of the 13-week study period, all samples were stored at –80 °C. Aliquots were made and DNA was isolated with the QIAamp DNA Stool Mini Kit. Isolated DNA was sequenced at the Beijing Genomics Institute (BGI).

Sequencing

We used 454 pyrosequencing to determine the bacterial composition of the fecal samples. Hyper-variable region V3 to V4 was selected using forward primer F515 (GTGCCAGCMGCCGCGG) and reverse primer: "E. coli 907-924" (CCGTCAATTCMTTTRAGT) to examine the bacterial composition.

We used QIIME [12], v1.7.0, to process the raw data files from the sequencer. The raw data files, sff files, were processed with the defaults of QIIME v1.7.0, however we did not trim the primers. Six out of 161 samples had fewer than 3000 reads and were excluded from the analysis. The average number of reads was 5862, with a maximum of 12,000 reads.

OTU picking

The operational taxonomic unit (OTU) formation was performed using the QIIME reference optimal picking, which uses UCLUST [13], version 1.2.22q, to perform the clustering. As a reference database, we used a primer-specific version of the full GreenGenes 13.5 database [14].

Using TaxMan [15], we created the primer-specific reference database, containing only reference entries that matched our selected primers. During this process we restricted the mismatches of the probes to the references to a maximum of 25 %. The 16S regions that were captured by our primers, including the primer sequences, were extracted from the full 16S sequences. For each of the reference clusters, we determined the overlapping part of the taxonomy of each of the reference reads in the clusters and used this overlapping part as the taxonomic label for the cluster. This is similar to the processes described in other studies [9, 15–18].

OTUs had to be supported by at least 100 reads and had to be identified in two samples; less abundant OTUs were excluded from the analysis.

Estimation of gene abundance and pathway activity

After filtering the OTUs, we used PICRUSt [19] to estimate the gene abundance and the PICRUSt output was then used in HUMAnN [20] to calculate the bacterial pathway activity. First, the reference database was clustered based on 97 % similarity to the reference sequence to better reflect the normal GreenGenes 97 % database required for PICRUSt. Three out of 1166 OTUs did not

contain a representative sequence in the GreenGenes 97 % set and were therefore excluded from the analysis. Since merging the reference database at 97 % similarity level led to merging of previously different clusters, for the pathway analysis we chose to permute the cluster representative names in the OTU-table 25 times; this was to be sure that our OTU picking strategy would not cause any problems in estimating the genes present in each micro-organism. Next, we ran PICRUSt on the 25 permuted tables and calculated the average gene abundance per sample. The average correlations between the permutations within a sample was higher than 0.97 (Pearson r). Hence, we averaged the PICRUSt output, which was then used to calculate the pathway activity in HUMAnN.

Changes in the gut microbiome or in gene abundance due to diet

To identify differentially abundant taxa, microbial biomarkers, and differences in pathway activity between the GFD and HD periods, we used QIIME and MaAsLin [21]. QIIME was used for the alpha-diversity analysis, principal coordinate analysis (PCoA) over unifrac distances, and visualization. In the MaAsLin analysis we corrected for ethnicity (defined as continent of birth) and gender. MaAsLin was used to search for differentially abundant taxonomic units to discriminate between the GFD and HD time points. Additionally, we tested for during transition from HD to GFD (T0–T4). MaAsLin uses a boosted, additive, general linear model to discriminate between groups of data.

In the MaAsLin analysis we did not test individual OTUs, but focused on the most detailed taxonomic label each OTU represented. Using the QIIMETOMAASLIN [22] tool, we aggregated the OTUs if the taxonomic label was identical and, if multiple OTUs represented a higher order taxa, we added this higher order taxa to the analysis. In this process, we went from 1166 OTUs to 114 separate taxonomic units that were included in our analysis. Using the same tool, QIIMETOMAASLIN, we normalized the microbial abundance using acrsin square root transformation. This transformation leads to the percentages being normally distributed.

In all our analyses we used the Q-value calculated using the R [23] Q-value package [24] to correct for multiple testing. The Q-value is the minimal false discovery rate at which a test may be called significant. We used a Q-value of 0.05 as a cutoff in our analyses.

Biomarkers

Six biomarkers related to gut health were measured in the "Dr. Stein & Colleagues" medical laboratory (Maastricht, the Netherlands). These biomarkers included: fecal calprotectin and a set of plasma cytokines as

Bonder *et al. Genome Medicine* (2016) 8:45

Page 4 of 11

markers for the immune system activation [25–27]; fecal human-β-defensin-2 as a marker for defense against invading microbes [28, 29]; fecal chromogranin A as a marker for neuro-endocrine system activation [30–32]; fecal short-chain fatty acids (SCFA) secretion as a marker for colonic metabolism [33]; and plasma citrulline as a measure for enterocyte mass [34, 35]. The plasma citrulline level and the panel of cytokines (IL-1β, IL-6, IL-8, IL-10, IL-12, and TNFα) were measured by high-performance liquid chromatography (HPLC) and electro-chemiluminescence immunoassay (ECLIA), respectively. In feces, we measured calprotectin and human-β-defensin-2 levels by enzyme-linked immunosorbent assay (ELISA), chromogranin A level by radioimmunoassay (RIA), and the short-chain fatty acids acetate, propionate, butyrate, valerate, and caproate by gas chromatography–mass spectrometry (GC-MS). All biomarker analyses were performed non-parametrically, with tie handling, because of the high number of samples with biomarker levels below the detection limit. We used the Wilcoxon test to compare the average biomarker levels between the diet periods and the Spearman correlation to search for relations between the microbiome or gene activity data and the biomarker levels.

Results

Food intake

We first investigated if a GFD had a significant effect on the daily intake of macronutrients by analyzing the GFD and HD food records from participants (Additional file 2: Table S1). Mean (SD) daily intakes of energy, protein, fat, and carbohydrate during GFD and HD are shown in Table 1. We observed slightly higher carbohydrate intake and a slightly lower fat intake on GFD; however, none of the differences in energy or macronutrient intake were significantly different. We therefore concluded that dietary macronutrient composition was not significantly changed by following a GFD.

Table 1 Mean and standard deviation (SD) of energy, protein, carbohydrates, and fat intake during the gluten-free diet (GFD) and habitual diet (HD). g = grams, en% = energy %

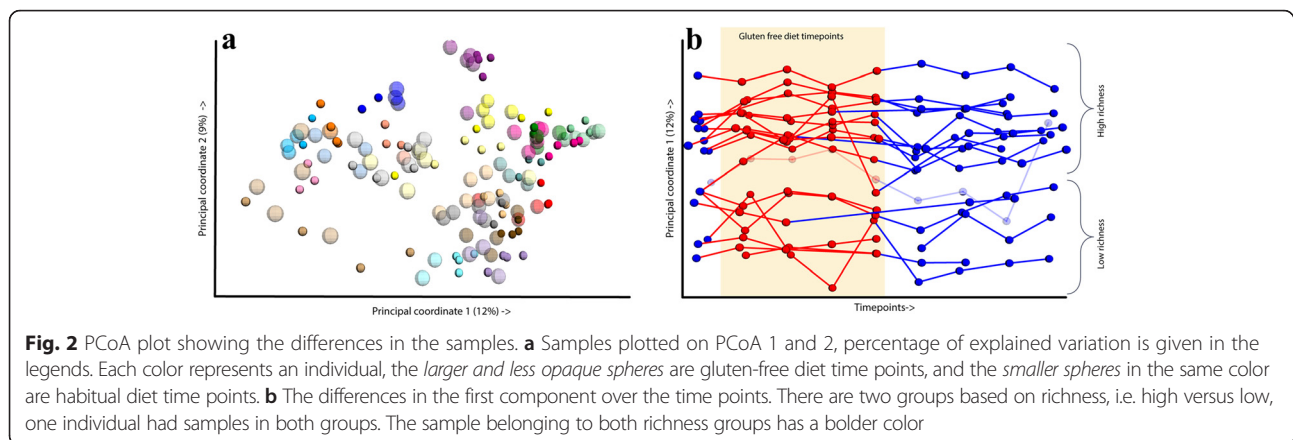| Nutrient	GFD (n = 12)		HD (n = 12)		
	Mean	SD	Mean	SD	p value
Energy (kcal)	1709.5	344.0	1811.5	433.9	0.243
Protein (g)	73.1	18.4	78.1	18.2	0.401
Protein (en%)	17.1		17.2		
Carbohydrates (g)	211.1	50.3	199.9	63.2	0.275
Carbohydrates (en%)	49.4		44.1		
Fat (g)	63.7	18.1	72.5	24.3	0.109
Fat (en%)	33.6		36.0		

Microbial differences due to diet

In total we used 155 fecal samples, originating from 21 individuals, for the microbiota analysis and we observed 114 different taxonomic units. We first checked if GFD influenced the number and proportion of bacteria in individual participants, for which we investigated differences in alpha diversity between the GFD and HD time points using several alpha diversity measures (Observed species, Shannon, Chao1, and Simpson indexes). We found no differences in the alpha diversity in any of these tests. Therefore, we concluded that a change in diet did not influence the bacterial diversity within a sample.

Next, we tested if there was any difference in the bacterial diversity related to variation in diet between participants (beta-diversity) by comparing the unweighted unifrac distance in sample groups. We observed a strong difference when comparing different time points from a single individual to all other individuals, regardless of diet type, Wilcoxon p value $<2.2 \times 10^{-16}$. When we compared the diet-induced differences within the same individual, we saw a small but significant change, Wilcoxon p value = 0.024, although the same diet time points were slightly more alike (Additional file 3: Figure S2).

In the PCoA analysis over the unweighted unifrac distance (Fig. 2a), we also saw that the main driver of the diversity is the inter-individual difference, with participants clustering together, both during and after the dietary intervention. In the first ten principal coordinates, which explain more than half of the total variation, we observed changes between the time points for individual participants, although there was no single component, or combination of components, capturing the difference between the GFD versus HD time points in the first ten components.

We therefore concluded that a GFD has a significant effect on the diversity between the groups, but that the inter-individual effect on the variation of the microbiome is stronger than the effect of diet.

We further investigated changes in beta-diversity in relation to the time points (Fig. 2b). When we plotted PCo1 versus the time points, we observed a separation into two groups. Since PCo1 describes the difference in alpha-diversity between samples, we concluded that this separation is based on richness. The richness separates all but one participant into either a clear high-richness or low-richness group (Fig. 2b). There is a significant difference in richness between the two groups, Wilcoxon p value = 0.0016, excluding the one participant who seems to be an intermediate. However, unlike the study by Le Chatelier et al. [36], we did not see any significant difference in stability, i.e. in variation in richness, between the low- and high-richness groups.

Bonder et al. Genome Medicine (2016) 8:45

Page 5 of 11

Fig. 2 PCoA plot showing the differences in the samples. **a** Samples plotted on PCoA 1 and 2, percentage of explained variation is given in the legends. Each color represents an individual, the *larger and less opaque spheres* are gluten-free diet time points, and the *smaller spheres* in the same color are habitual diet time points. **b** The differences in the first component over the time points. There are two groups based on richness, i.e. high versus low, one individual had samples in both groups. The sample belonging to both richness groups has a bolder color

Differentially abundant taxa

When comparing the HD and GFD time points, corrected for age and ethnicity in MaAsLin, we observed eight significant microbial changes (Fig. 3 and Table 2). The strongest association was found to the family *Veillonellaceae*, of which the abundance in the gut dropped significantly on a GFD ($p = 2.81 \times 10^{-05}$, q = 0.003) (Fig. 3b and Additional file 4: Figure S3). Other species that decreased on a GFD included *Ruminococcus bromii* ($p = 0.0003$, q = 0.01) and *Roseburia faecis* ($p = 0.002$, q = 0.03). While families *Victivallaceae* ($p = 0.0002$, q = 0.01), *Clostridiaceae* ($p = 0.0006$, q = 0.015), and *Coriobacteriaceae* ($p = 0.003$, q = 0.035), order *ML615J-28* ($p = 0.001$, q = 0.027), and genus *Slackia* ($p = 0.002$, q = 0.01) increased in abundance on a GFD.

Next, we tested for trends during the diet change; however, we did not observe a time-dependent change in the microbiome composition. Since we observed two different groups based on richness in the PCoA analysis, we tested for different reactions to the change in diet in the high-richness- and low-richness groups. However, no significant associations were found in this analysis.

Since six out of the 28 participants smoked, we tested for overlap between smoke-associated bacteria and diet-related bacteria. We did not find any overlap; Additional file 5: Table S2 shows the bacteria associated with smoking.

Imputation of bacterial function

Next to the taxonomic associations, we also aimed to study differences in pathway composition in relation to GFD. We applied PICRUSt and HUMAnN for pathway annotation, as described in Methods. In total, 161 pathways and 100 modules were predicted, all of the pathways and modules were found in at least 1 % of the samples.

We used MaAsLin to identify differences in the pathway composition and conducted the same tests – GFD versus HD and the time-series test – as for the microbial
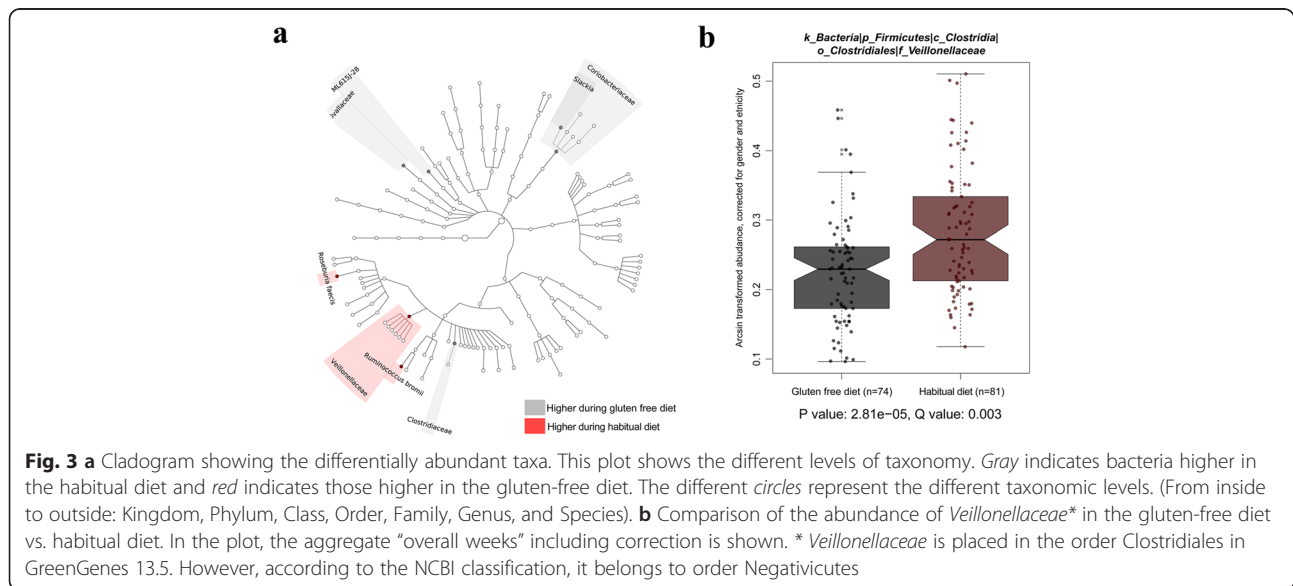

Fig. 3 a Cladogram showing the differentially abundant taxa. This plot shows the different levels of taxonomy. *Gray* indicates bacteria higher in the habitual diet and *red* indicates those higher in the gluten-free diet. The different *circles* represent the different taxonomic levels. (From inside to outside: Kingdom, Phylum, Class, Order, Family, Genus, and Species). **b** Comparison of the abundance of *Veillonellaceae** in the gluten-free diet vs. habitual diet. In the plot, the aggregate "overall weeks" including correction is shown. * *Veillonellaceae* is placed in the order Clostridiales in GreenGenes 13.5. However, according to the NCBI classification, it belongs to order Negativicutes

Bonder *et al. Genome Medicine* (2016) 8:45

Page 6 of 11

Table 2 GFD-induced changes in taxonomic composition

Taxonomic unit	Coefficient	N.not.0/N	p value	Q-value
p_Firmicutes|c_Clostridia|o_Clostridiales|f_Veillonellaceae[a]	0.0424	155/155	2.81×10^{-5}	0.0030
p_Lentisphaerae|c_Lentisphaeria|o_Victivallales|f_Victivallaceae	−0.0093	89/155	2.30×10^{-4}	0.0105
p_Firmicutes|c_Clostridia|o_Clostridiales|f_Ruminococcaceae|g_Ruminococcus|s_bromii	0.0151	99/155	2.94×10^{-4}	0.0105
p_Firmicutes|c_Clostridia|o_Clostridiales|f_Clostridiaceae	−0.0121	150/155	5.69×10^{-4}	0.0152
p_Tenericutes|c_RF3|o_ML615J-28	−0.0095	82/155	1.30×10^{-3}	0.0277
p_Firmicutes|c_Clostridia|o_Clostridiales|f_Lachnospiraceae|g_Roseburia|s_faecis	0.0065	100/155	1.88×10^{-3}	0.0326
p_Actinobacteria|c_Coriobacteriia|o_Coriobacteriales|f_Coriobacteriaceae|g_Slackia	−0.0044	43/155	2.14×10^{-3}	0.0326
p_Actinobacteria|c_Coriobacteriia|o_Coriobacteriales|f_Coriobacteriaceae	−0.0137	155/155	2.67×10^{-3}	0.0357

A positive coefficient means more of the microbe was present during the habitual diet, while a negative coefficient means less of the microbe was present during the habitual diet. All associations were to the kingdom bacteria, for readability the kingdom label is not presented. [a]*Veillonellaceae* is placed in the order Clostridiales in GreenGenes 13.5. However, according to the NCBI classification, it belongs to order Negativicutes

composition. The data were again corrected for age and ethnicity. We observed that 19 KEGG pathways and two KEGG modules (Table 3) were different in abundance between GFD and HD. We did not observe associations related to the transition from GFD to HD (T0–T4). Four out of five top associations, all with a Q-value <0.0003, are related to metabolism changes: tryptophan metabolism, butyrate metabolism (Fig. 4a), fatty acid metabolism, and seleno-compound metabolism.
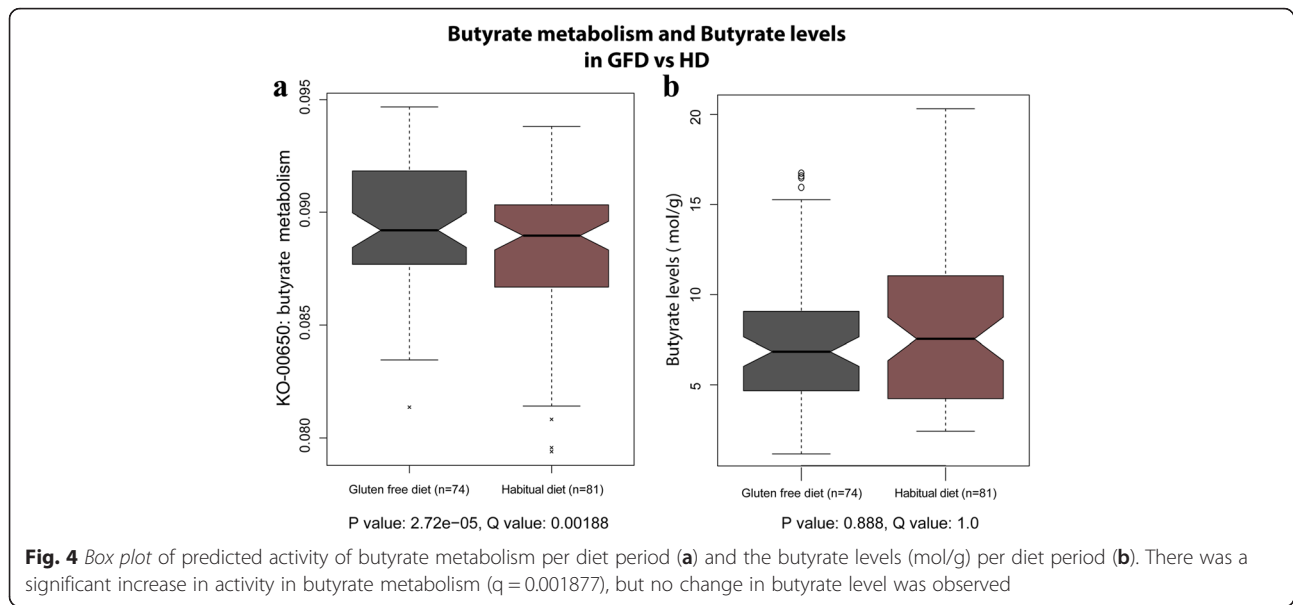
Biomarkers in relation to diet changes
Biomarkers related to GFD versus HD

We measured four biomarkers in feces: calprotectin, human-β-defensin-2, chromogranin A, and a set of five short-chain fatty acids (acetate, propionate, butyrate, valerate, and caproate). In addition, we measured citrulline levels and a panel of cytokines (IL-1β, IL-6, IL-8, IL-10, IL-12, and TNFα) in blood. The Wilcoxon test was used to test biomarker level differences between the average

Table 3 GFD-induced changes in pathway and module activity

Feature	Coefficient	N.not.0/N	p value	Q-value
KO00380: Tryptophan metabolism	−0.0011	155/155	2.45×10^{-5}	0.002
KO00650: Butyrate metabolism	−0.0014	155/155	2.72×10^{-5}	0.002
KO00071: Fatty acid metabolism	−0.0011	155/155	4.74×10^{-5}	0.002
KO00450: Selenocompound metabolism	0.0009	155/155	9.23×10^{-5}	0.003
KO00630: Glyoxylate and dicarboxylate metabolism	−0.0010	155/155	2.53×10^{-4}	0.007
KO00520 Amino sugar and nucleotide sugar metabolism	0.0009	155/155	2.83×10^{-4}	0.007
M00064: ADP-L-glycero-D-manno-heptose biosynthesis	0.0066	155/155	4.12×10^{-4}	0.023
KO00643: Styrene degradation	−0.0013	155/155	4.29×10^{-4}	0.008
M00077: Chondroitin sulphate degradation Chondroitin sulphate degradation	−0.0037	76/155	5.81×10^{-4}	0.023
KO00760: Nicotinate and nicotinamide metabolism	0.0008	155/155	6.79×10^{-4}	0.012
KO00620: Pyruvate metabolism	−0.0012	155/155	0.002	0.023
KO00253: Tetracycline biosynthesis	−0.0027	155/155	0.002	0.024
KO00471: D-Glutamine and D-glutamate metabolism	0.0012	155/155	0.002	0.024
KO04122: Sulphur relay system	−0.0020	155/155	0.002	0.024
KO00633: Nitrotoluene degradation	−0.0022	155/155	0.002	0.024
KO00072: Synthesis and degradation of ketone bodies	−0.0020	155/155	0.003	0.028
KO00310: Lysine degradation	−0.0007	155/155	0.003	0.031
KO00624: Polycyclic aromatic hydrocarbon degradation	0.0006	155/155	0.005	0.043
KO00561: Glycerolipid metabolism	−0.0012	155/155	0.005	0.043
KO00680: Methane metabolism	−0.0006	155/155	0.006	0.047
KO00550: Peptidoglycan biosynthesis	0.0011	155/155	0.007	0.047

A positive coefficient means more activity of the pathway/module during the habitual diet, while a negative coefficient means less activity of the pathway/module during the habitual diet

Bonder *et al. Genome Medicine* (2016) 8:45

Page 7 of 11

Fig. 4 *Box plot* of predicted activity of butyrate metabolism per diet period (**a**) and the butyrate levels (mol/g) per diet period (**b**). There was a significant increase in activity in butyrate metabolism (q = 0.001877), but no change in butyrate level was observed

values and the GFD and HD period values. We saw no significant change in biomarker levels in relation to GFD (Table 4A and B).

Correlations between fecal biomarkers and microbiome

We correlated the fecal biomarker levels to the microbiome composition as well as to the microbiome predicted pathways and modules. After multiple testing correction, we observed many statistically significant correlations between the levels of biomarkers and microbiome/

pathway abundances; the absolute correlation, Spearman Rho, was between 0.14 and 0.59. An expected observation was the correlation of the butyrate pathway activity to the butyrate biomarker, as we had previously observed a significant correlation between the predicted butyrate pathway activity and diet change (Table 3). When correlating the actual butyrate measurements with the predicted activity of the butyrate metabolism, we observed a low but significant correlation of −0.269 ($p = 0.0009$, q = 0.0012, Additional file 6: Figure S4). However, there was no

Table 4 Median and 25 %/75 % quantiles of the measured biomarkers

	Habitual diet	Gluten-free diet	Wilcoxon test *p* value
A) Plasma			
Citrullin (mol/L)	45.60 (38.15–51.50)	48.00 (36.35–56.85)	0.9328
IL 1 Beta (g/L)	1.60 (0.68–2.10)	1.23 (0.79–1.68)	0.8870
IL 6 (g/L)	BDL (BDL–1.60)	BDL (BDL–0.38)	0.1240
IL 8 (g/L)	6.04 (2.89–12.61)	5.41 (3.34–11.19)	0.9030
IL 10 (g/L)	0.83 (0.74–1.01)	0.83 (0.74–0.97)	0.9322
IL 12P70 (g/L)	1.53 (0.95–1.78)	1.53 (0.95–2.11)	0.2131
TNF Alpha (g/L)	0.56 (BDL–4.33)	BDL (BDL–5.13)	0.9761
B) Feces			
Chromogranin A (nmol/g)	10.85 (7.69–23.09)	11.44 (7.37–27.18)	0.8128
Beta Defensin 2 (ng/g)	24.90 (18.78–35.03)	26.10 (20.03–46.90)	0.5256
Calprotectin (g/g)	21.55 (BDL–42.88)	13.05 (BDL–31.28)	0.0528
Acetate (mol/g)	24.37 (17.35–34.34)	23.61 (18.58–35.12)	0.8651
Propionate (mol/g)	7.55 (4.24–10.98)	6.84 (4.67–9.07)	0.6986
Butyrate (mol/g)	6.86 (3.53–10.63)	6.48 (4.27–10.40)	0.8882
Valerat (mol/g)	1.09 (0.74–1.76)	1.24 (0.79–1.70)	0.6824
Caproat (mol/g)	0.28 (0.05–0.85)	0.21 (0.04–0.66)	0.2488

None of the differences were statistically significant. BDL = below detection limit

Bonder *et al. Genome Medicine* (2016) 8:45

Page 8 of 11

significant difference in butyrate levels in the two diet periods (Fig. 4b and Table 4). Another interesting correlation was found between the predicted pyruvate metabolism pathway and the levels of propionate (mol/g), since propionate can be oxidized to pyruvate [37], for which we observed a correlation of -0.54 ($p = 9.44 \times 10^{-13}$, q = 1.48×10^{-10}, Additional file 7: Figure S5). A complete list of the significant correlations between the fecal biomarkers and the microbiome compositions, the predicted KEGG pathway activity scores, and predicted activity of KEGG modules can be found in Additional file 8: Tables S3, Additional file 9: Table S4, and Additional file 10: Table S5.

Discussion

We investigated the role of a four-week GFD on microbiome composition in healthy individuals and identified moderate but significant changes in their microbiome compositions and even stronger effects on the imputed activity levels of bacterial pathways.

On a taxonomic level we identified eight bacteria that change significantly in abundance on GFD: *Veillonellaceae*, *Ruminococcus bromii*, and *Roseburia faecis* decreased on GFD, and *Victivallaceae*, *Clostridiaceae*, *ML615J-28*, *Slackia*, and *Coriobacteriaceae* increased on GFD. The strongest effect was seen in the decrease of *Veillonellaceae* during GFD, Gram-negative bacteria known for lactate fermentation. This is the first time that the *Veillonellaceae* family has been associated to a dietary intervention, but it was recently shown to be decreased in autistic patients [38]. Remarkably, the patients in that study were more often on a GFD (9/10) than the control group (5/10). Our findings suggest that GFD, rather than autism, can be the cause of a lower abundance of *Veillonellaceae* in these patients, thus highlighting the importance of including dietary information in analyses of microbiota in relation to diseases. *Veillonellaceae* is considered to be a pro-inflammatory family of bacteria; an increase in *Veillonellaceae* abundance was consistently reported in IBD, IBS, and cirrhosis patients [39–41]. It is conceivable that a decrease in *Veillonellaceae* abundance might be one of the mediators of the GFD's beneficial effect observed in patients with IBS and gluten-related disorders.

Several of the associated bacteria have been previously linked to diet changes and starch metabolism. In particular, *Ruminococcus bromii* is important for the degradation of resistant starch in the human colon [42] and is increased when on a resistant starch diet [43]. It is also known that degradation of cellulose by *Ruminococcus* results in the production of SCFA and hydrogen gas [44]; a decrease in abundance of *Ruminococcus* and its fermentation products might explain the beneficial effect of a GFD that is experienced by some IBS patients as

previously reported by Aziz et al. [45]. Both *Ruminococcus bromii* and *Roseburia faecis* were recently reported to be influenced by switching from a vegetarian to a meat-containing diet [9]. It is likely that changes in these bacteria observed in relation to GFD are the consequences of the different starch composition of a GFD versus HD. Moreover, stool consistency could influence the results of microbiome composition [46]; unfortunately, data on stool composition were not collected in our study.

The five bacteria for which we found an increased abundance on GFD are less well characterized although the *Slackia* genus, its family *Coriobacteriaceae*, and the family *Clostridiaceae* have been previously linked to gastrointestinal diseases in humans – inflammatory bowel disease, celiac disease, and colorectal cancer [47–49]. The *Victivallaceae* family and ML615J-28 order have not been previously associated to diet change or phenotypic change in human. However, in general, it could be hypothesized that these bacteria benefit from a change in available substrates as a result from the change in diet, which could in turn result in altered metabolite production and related gastrointestinal complaints.

In this study we found a stronger effect of diet on the imputed KEGG pathways than on the taxonomic level. So, although the changes in the overall microbiome were moderate, there were more profound effects on the pathway activities of the microbiome.

The strength of our study lies in our analysis of the microbiome at multiple time points for the same individuals. We identified that the inter-individual variability is the strongest determinant of sample variability, suggesting that in healthy individuals the gut microbiome is stable, even with short-term changes in the habitual diet. We did not observe differences in the downstream effect of GFD in relation to high or low richness, which contradicts previous observations [50]. The study by David et al. [9] identified a profound effect of short-term diet change from a vegetarian to an animal-based diet and vice versa. This profound short-term diet effect was not observed in our study when changing from a gluten-containing to a gluten-free diet. Induced by the diet change, David et al. [9] found significant differences in macronutrient intake between meat-based and plant-based diet, whereas macronutrient intake in this study was not changed during the diets. These results suggest that changing the main energy source (meat vs. plant) has a more profound effect on the microbiome than changing the carbohydrate source (gluten). Although De Palma et al. [8] did observe a reduction in polysaccharide intake for GFD in healthy individuals, we were unable to reproduce their finding because we could not distinguish between different classes of carbohydrates in our dataset as the food composition data on GFD

Bonder *et al. Genome Medicine* (2016) 8:45

Page 9 of 11

foods lacked this information. Further, it is possible that changes in nutritional intake other than those driven by gluten exclusion might influence microbiome changes.

For our selection of blood and stool biomarkers, we observed no significant associations with the diet change. All the selected biomarkers are markers of inflammation or metabolic changes and remained in the normal range in all our participants, with a high proportion of the values of blood inflammatory markers being below the detection limit. Overall, we conclude that a GFD and its downstream effects on the microbiome do not cause major inflammatory or metabolic changes in gut function in healthy participants. However, the lower abundance of *Veillonellaceae*, the pro-inflammatory bacterium linked to Crohn's disease and other gut disease phenotypes, suggests a reduction in gut inflammatory state. This change in bacterial composition might be linked with a beneficial effect of GFD for patients with gut disorders such as gluten-related disorders and/or IBS.

Conclusions

We have identified eight taxa and 21 bacterial pathways associated with a change from a habitual diet to a GFD in healthy individuals. We conclude that the effect of gluten intake on the microbiota is less pronounced than that seen for a shift from a meat-based diet to a vegetarian diet (or vice versa). However, a GFD diet clearly influences the abundance of several species, in particular those involved specifically in carbohydrate and starch metabolism. Our study illustrates that variations in diet could confound the results of microbiome analysis in relation to disease phenotypes, so dietary variations should be carefully considered and reported in such studies. The short-term GFD did not influence the levels of inflammatory gut biomarkers in healthy individuals. Further research is needed to assess the impact of a GFD on inflammatory and metabolic changes in gut function in individuals with gastrointestinal conditions such as IBS and gluten-related disorders.

Ethics approval and consent to participate

This GFD study followed the sampling protocol of the LifeLines-DEEP study, which was approved by the ethics committee of the University Medical Centre Groningen and conform the Declaration of Helsinki, document no. METC UMCG LLDEEP: M12.113965. All participants signed their informed consent prior to study enrolment.

Availability of data and materials

The supporting data are available to researchers in the European Nucleotide Archive, under study accession number PRJEB13219 (http://www.ebi.ac.uk/ena/data/view/PRJEB13219).

Additional files

Additional file 1: Figure S1. Baseline characteristics of the GFD study group. (TIF 1068 kb)

Additional file 2: Table S1. Results macronutrient intake per participant. (XLSX 13 kb)

Additional file 3: Figure S2. Unweighted unifrac distances when comparing inter-individual vs intra individual distances. In group 1 the intra-individual differences are shown regardless of diet. Group 2 shows the intra-sample differences are shown within the same diet. Group 3 shows the intra-individual differences are shown between the two diet groups. In group 4 the inter-individual differences are shown regardless of diet. Group 5 shows the inter-sample differences are shown within the same diet. Group 6 shows the inter-individual differences are shown between the two diet groups. The main difference is the intra- vs. inter-individual difference. Also the same diet points in the samples are slightly closer to each other. However, we do not see such a phenomenon for group 5 vs. group 6. (TIF 1862 kb)

Additional file 4: Figure S3. Abundance of Veillonellaceae family in the GFD participants. In all but four participants we see a clear trend of higher levels of Veillonellaceae on the habitual diet. The rightmost samples do not show this phenomenon. (TIF 4036 kb)

Additional file 5: Table S2. Relation of smoking and microbiome composition. (XLSX 9 kb)

Additional file 6: Figure S4. Measured butyrate levels vs. the predicted activity of butyrate metabolism. (TIF 6 kb)

Additional file 7: Figure S5. Measured propionate levels vs. the predicted activity of pyruvate metabolism. (TIF 646 kb)

Additional file 8: Table S3. Correlation of bacteria and levels of fecal biomarkers. (PDF 139 kb)

Additional file 9: Table S4. Correlation of predicted HUMAnN pathway activity and levels of fecal biomarkers. (PDF 228 kb)

Additional file 10: Table S5. Correlation of predicted HUMAnN module activity and levels of fecal biomarkers. (PDF 143 kb)

Abbreviations

BGI: Beijing Genomics Institute; ECLIA: electro-chemiluminescence immunoassay; EDTA: ethylenediaminetetraacetic acid; ELISA: enzyme-linked immunosorbent essay; FISH: fluorescence in situ hybridization; GC-MS: gas chromatography–mass spectrometry; GFD: gluten-free diet; HD: habitual diet; HPLC: high performance liquid chromatography; IBS: irritable bowel syndrome; KEGG: Kyoto encyclopedia of genes and genomes; OTU: operational taxonomic unit; PCoA: principal coordinate analysis; qPCR: quantitative real-time polymerase chain reaction; RIA: radioimmunoassay; SCFA: short chain fatty acids; SD: standard deviation.

Bonder *et al. Genome Medicine* (2016) 8:45

Page 10 of 11

Author details
[1]Department of Genetics, University of Groningen, University Medical Centre Groningen, Groningen, The Netherlands. [2]Top Institute Food and Nutrition, Wageningen, The Netherlands. [3]BGI, Shenzhen 518083, China. [4]Wellcome Trust Sanger Institute, Hinxton, Cambridge CB10 1SA, UK. [5]Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA. [6]Department of Computer Science, Aalto University School of Science, Espoo 02150, Finland.

References
1. Sollid LM. Coeliac disease: dissecting a complex inflammatory disorder. Nat Rev Immunol. 2002;2:647–55.
2. Sapone A, Bai JC, Ciacci C, Dolinsek J, Green PH, Hadjivassiliou M, et al. Spectrum of gluten-related disorders: consensus on new nomenclature and classification. BMC Med. 2012;10:13.
3. Catassi C, Bai JC, Bonaz B, Bouma G, Calabrò A, Carroccio A, et al. Non-celiac gluten sensitivity: the new frontier of gluten related disorders. Nutrients. 2013;5:3839–53.
4. Vazquez-Roque MI, Camilleri M, Smyrk T, Murray JA, Marietta E, O'Neill J, et al. A controlled trial of gluten-free diet in patients with irritable bowel syndrome-diarrhea: effects on bowel frequency and intestinal function. Gastroenterology. 2013;144:903–911.e3.
5. Collado MC, Donat E, Ribes-Koninckx C, Calabuig M, Sanz Y. Specific duodenal and faecal bacterial groups associated with paediatric coeliac disease. J Clin Pathol. 2009;62:264–9.
6. Di Cagno R, Rizzello CG, Gagliardi F, Ricciuti P, Ndagijimana M, Francavilla R, et al. Different fecal microbiotas and volatile organic compounds in treated and untreated children with celiac disease. Appl Environ Microbiol. 2009;75:3963–71.
7. Nistal E, Caminero A, Vivas S, de Morales JMR, de Miera LES, Rodríguez-Aparicio LB, et al. Differences in faecal bacteria populations and faecal bacteria metabolism in healthy adults and celiac disease patients. Biochimie. 2012;94:1724–9.
8. De Palma G, Nadal I, Collado MC, Sanz Y. Effects of a gluten-free diet on gut microbiota and immune function in healthy adult human subjects. Br J Nutr. 2009;102:1154–60.
9. David LA, Maurice CF, Carmody RN, Gootenberg DB, Button JE, Wolfe BE, et al. Diet rapidly and reproducibly alters the human gut microbiome. Nature. 2014;505:559–63.
10. Tigchelaar EF, Zhernakova A, Dekens JAM, Hermes G, Baranska A, Mujagic Z, et al. An Introduction to LifeLines DEEP: Study Design and Baseline Characteristics. Cold Spring Harbor Labs J. 2014.
11. Baranska A, Tigchelaar E, Smolinska A, Dallinga JW, Moonen EJC, Dekens JAM, et al. Profile of volatile organic compounds in exhaled breath changes as a result of gluten-free diet. J Breath Res. 2013;7:037104.
12. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK. QIIME allows analysis of high-throughput community sequencing data. Nat Methods. 2010;7:335–6.
13. Edgar RC. Search and clustering orders of magnitude faster than BLAST. Bioinformatics. 2010;26:2460–1.
14. DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, et al. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. Appl Environ Microbiol. 2006;72:5069–72.
15. Brandt BW, Bonder MJ, Huse SM, Zaura E. TaxMan: a server to trim rRNA reference databases and inspect taxonomic coverage. Nucleic Acids Res. 2012;40(Web Server issue)):W82–7.
16. Bonder MJ, Abeln S, Zaura E, Brandt BW. Comparing clustering and pre-processing in taxonomy analysis. Bioinformatics. 2012;28:2891–7.
17. May A, Abeln S, Crielaard W, Heringa J, Brandt BW. Unraveling the outcome of 16S rDNA-based taxonomy analysis through mock data and simulations. Bioinformatics. 2014;30:1530–8.
18. Ding T, Schloss PD. Dynamics and associations of microbial community types across the human body. Nature. 2014;509:357–60.
19. Langille MGI, Zaneveld J, Caporaso JG, McDonald D, Knights D, Reyes J, et al. Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. Nat Biotechnol. 2013;31:814–21.
20. Abubucker S, Segata N, Goll J, Schubert AM, Izard J, Cantarel BL, et al. Metabolic reconstruction for metagenomic data and its application to the human microbiome. PLoS Comput Biol. 2012;8:e1002358.
21. Tickle T, Waldron L, Yiren Lu, Huttenhower C. Multivariate association of microbial communities with rich metadata in high-dimensional studies. (In progress)
22. Tickle T. QiimeToMaAsLin. 2013.
23. R Development Core Team RFFSC. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna: R Development; 2008.
24. Dabney A, Storey JD, qvalue Tutorial PRS, Install R, Packages IB, Packages UB, Vignettes P, Cloud AMI, Team C. Q-value estimation for false discovery rate control. Medicine. 2004:539–548
25. Tibble JA, Sigthorsson G, Foster R, Scott D, Fagerhol MK, Roseth A, et al. High prevalence of NSAID enteropathy as shown by a simple faecal test. Gut. 1999;45:362–6.
26. Joshi S, Lewis SJ, Creanor S, Ayling RM. Age-related faecal calprotectin, lactoferrin and tumour M2-PK concentrations in healthy volunteers. Ann Clin Biochem. 2010;47(Pt 3):259–63.
27. Maheshwari A, Lu W, Lacson A, Barleycorn AA, Nolan S, Christensen RD, et al. Effects of interleukin-8 on the developing human intestine. Cytokine. 2002;20:256–67.
28. Harder J, Bartels J, Christophers E, Schröder JM. A peptide antibiotic from human skin. Nature. 1997;387:861.
29. Langhorst J, Junge A, Rueffer A, Wehkamp J, Foell D, Michalsen A, et al. Elevated human beta-defensin-2 levels indicate an activation of the innate immune system in patients with irritable bowel syndrome. Am J Gastroenterol. 2009;104:404–10.
30. El-Salhy M, Lomholt-Beck B, Hausken T. Chromogranin A as a possible tool in the diagnosis of irritable bowel syndrome. Scand J Gastroenterol. 2010;45:1435–9.
31. Sidhu R, Drew K, McAlindon ME, Lobo AJ, Sanders DS. Elevated serum chromogranin A in irritable bowel syndrome (IBS) and inflammatory bowel disease (IBD): a shared model for pathogenesis? Inflamm Bowel Dis. 2010;16:361.
32. Ohman L, Stridsberg M, Isaksson S, Jerlstad P, Simrén M. Altered levels of fecal chromogranins and secretogranins in IBS: relevance for pathophysiology and symptoms? Am J Gastroenterol. 2012;107:440–7.
33. Hamer HM, Jonkers D, Venema K, Vanhoutvin S, Troost FJ, Brummer R-J. Review article: the role of butyrate on colonic function. Aliment Pharmacol Ther. 2008;27:104–19.
34. Windmueller HG, Spaeth AE. Source and fate of circulating citrulline. Am J Physiol. 1981;241:E473–80.
35. Crenn P, Messing B, Cynober L. Citrulline as a biomarker of intestinal failure due to enterocyte mass reduction. Clin Nutr. 2008;27:328–39.
36. Le Chatelier E, Nielsen T, Qin J, Prifti E, Hildebrand F, Falony G, et al. Richness of human gut microbiome correlates with metabolic markers. Nature. 2013;500:541–6.
37. Brock M, Maerker C, Schütz A, Völker U, Buckel W. Oxidation of propionate to pyruvate in Escherichia coli: Involvement of methylcitrate dehydratase and aconitase. Eur J Biochem. 2002;269:6184–94.
38. Kang D-W, Park JG, Ilhan ZE, Wallstrom G, LaBaer J, Adams JB, et al. Reduced incidence of prevotella and other fermenters in intestinal microflora of autistic children. PLoS One. 2013;8:e68322.
39. Gevers D, Kugathasan S, Denson LA, Vázquez-Baeza Y, Van Treuren W, Ren B, et al. The treatment-naive microbiome in new-onset Crohn's disease. Cell Host Microbe. 2014;15:382–92.
40. Haberman Y, Tickle TL, Dexheimer PJ, Kim M, Tang D, Karns R, et al. Pediatric Crohn disease patients exhibit specific ileal transcriptome and microbiome signature. J Clin Invest. 2014;124:3617–33.
41. Shukla R, Ghoshal U, Dhole TN, Ghoshal UC. Fecal microbiota in patients with irritable bowel syndrome compared with healthy controls using real-time polymerase chain reaction: an evidence of dysbiosis. Dig Dis Sci. 2015;60:2953–62.
42. Ze X, Duncan SH, Louis P, Flint HJ. Ruminococcus bromii is a keystone species for the degradation of resistant starch in the human colon. ISME J. 2012;6:1535–43.
43. Walker AW, Ince J, Duncan SH, Webster LM, Holtrop G, Ze X, et al. Dominant and diet-responsive groups of bacteria within the human colonic microbiota. ISME J. 2011;5:220–30.
44. Rajilić-Stojanović M. Function of the microbiota. Best Pract Res Clin Gastroenterol. 2013;27:5–16.
45. Aziz I, Trott N, Briggs R, North JR, Hadjivassiliou M, Sanders DS. Efficacy of a gluten-free diet in subjects with irritable bowel syndrome-diarrhea

Bonder *et al. Genome Medicine* (2016) 8:45

Page 11 of 11

unaware of their HLA-DQ2/8 Genotype. Clin Gastroenterol Hepatol. 2015. doi:10.1016/j.cgh.2015.12.031 [Epub ahead of print].

46. Tigchelaar EF, Bonder MJ, Jankipersadsing SA, Fu J, Wijmenga C, Zhernakova A. Gut microbiota composition associated with stool consistency. Gut. 2016; 65:540–2.

47. Maukonen J, Kolho K-L, Paasela M, Honkanen J, Klemetti P, Vaarala O, et al. Altered fecal microbiota in paediatric inflammatory bowel disease. J Crohns Colitis. 2015;9:1088–95.

48. Chen W, Liu F, Ling Z, Tong X, Xiang C. Human intestinal lumen and mucosa-associated microbiota in patients with colorectal cancer. PLoS One. 2012;7:e39743.

49. Olivares M, Neef A, Castillejo G, De Palma G, Varea V, Capilla A, et al. The HLA-DQ2 genotype selects for early intestinal microbiota composition in infants at high risk of developing coeliac disease. Gut. 2015;64:406–17.

50. Fang S, Evans RM. Microbiology: Wealth management in the gut. Nature. 2013;500:538–9.